# Cost-efficient service provision in neurorehabilitation: defining needs, costs and outcomes for people with long term neurological conditions

*Prepared on behalf of the UK Rehabilitation Outcomes Collaborative (UKROC) steering group by:*


**Lynne Turner-Stokes,**

**Roxana Vanderstay,**

**Kathy Eagar,**

**Robert Dredge,**

**Richard Siegert**


**This report was submitted to the NIHR for review on 29.7.2015.**

**It is currently still under review by the NIHR team, but may be read as a DRAFT report**

# Cost-efficient service provision in neurorehabilitation: defining needs, costs and outcomes for people with long term neurological conditions


## Lynne Turner-Stokes, Roxana Vanderstay, Kathy Eagar, Robert Dredge, Richard Siegert


### *A report prepared on behalf of the UK Rehabilitation Outcomes Collaborative (UKROC) steering group*


**With assistance from:**
**Diana Jackson[1], Keith Sephton[2], Alan Bill[2], Heather Williams[2], Frances Simmonds[5], Tara Stevermeuer[5], Paul Bassett[11], Ann and Steve Harris and the British Society of Rehabilitation Medicine**

### *UKROC Programme Steering Group Membership:*

**Lynne Turner-Stokes[1,2] , Richard Siegert[3], Roxana Vanderstay[1], Robert Dredge[4], Kathy Eagar[5], Derick Wade[6], Paul McCrone[7], Charles Wolfe[8], Virginia Jordan[9], Kellie Blane[10], Audrey Malik[2].**

1. Department of Palliative Care, Policy and Rehabilitation, Faculty of Life Sciences and Medicine, King's College London, London, UK.
2. Regional Rehabilitation Unit, North West London Hospitals Trust, Northwick Park Hospital, Harrow, Middlesex, UK.
3. School of Public Health & Psychosocial Studies and School of Rehabilitation & Occupation Studies, Auckland University of Technology, Auckland, New Zealand.
4. Centre for Health Planning Management, Keele University, Keele, UK.
5. Australian Health Services Research Institute (AHSRI), University of Wollongong, Wollongong, Australia.
6. Oxford Centre for Enablement, Nuffield Orthopaedic Centre and King's College London, UK.
7. Centre for the Economics of Physical and Mental Health, Institute of Psychiatry, King's College London, London, UK.
8. Department of Primary Care and Public Health Sciences, Division of Health and Social Care Research, Faculty of Life Sciences and Medicine, King's College London, London, UK.
9. The Health and Social Care Information Centre, London, UK.
10. NHS England, Victoria, London, UK.
11. Statsconsultancy, London, UK.

# Acknowledgements

The authors gratefully acknowledge the support given from the very many people who have contributed to this programme of work. They are too many to name individually, but individuals who have made a critical contribution are highlighted in the 'list of contributors'.

First and foremost we would like to thank the patients and clinical teams who contributed the data, and also the managers and service commissioners who have provided both data and helpful feedback.
We are particularly indebted to all of the patients and family members who have contributed to the various PPI groups, or provided feedback over the course of the programme.

We are grateful to members of the following groups / committees who have contributed their expertise constructive criticism and support:

- British Society of Rehabilitation Medicine Executive and Research and Clinical Standards Committees.
- The HSCIC Expert Reference Panel for Rehabilitation.
- The NHSE Clinical Reference Group for Specialist Rehabilitation.

We are also grateful to the various teams in Australia and New Zealand who have worked with us along the way, contributing their experience, enthusiasm and often data as well.

We would like in particular to recognise the contribution of Professor Bipin Bhakta to the planning and development of this programme. He sadly passed away before being able to see the outputs, which we believe he would have approved of.

# List of key contributors

| Individual | Affiliation | Contribution / expertise |
|---|---|---|
| **UKROC Programme Steering Group** | | |
| **Prof Lynne Turner-Stokes** | King's College London and Northwick Park Hospital | Principal Applicant, Director of UKROC. Lead Author for report – reposible for drafting |
| **Prof Richard J Siegert** | Auckland University of Technology and King's College London | Lead Psychometrician |
| **Dr Roxana Vanderstay** | King's College London | Database analysis |
| **Prof Kath Eager** | University of Wollongong | Director Australian Health Services Research Institute and Lead Collaborator for Australia |
| **Dr Robert Dredge** | Keele University | Lead for NHS Accountancy |
| **Co-applicants - current** | | |
| **Prof Derick Wade** | Oxford Centre for Enablement | Clinician and expertise in outcome measurement |
| **Prof Paul McCrone** | Institute of Psychiatry, King's College London | Health Economist |
| **Prof Charles Wolfe** | King's College London | Health Service Research and Registry Development |
| **Virginina Jordan** | Health and Social Care Information Centre | Casemix development |
| **Kellie Blane** | NHS England | Specialised rehabilitation services commissioning |
| **Co-applicants past** | | |
| **Barbara Howe** | NHS England | Expertise on national commissioining |
| **Stephen Sutch** | John Hopkins University | Casemix analysis |
| **Stephen Cole** | Health and Social Care Information Centre | Casemix development |
| **Gareth Dear** | | |
| **Rachel McIlroy** | | |
| **Sue Walker** | Northwick Park Hospital | NHS Manager |
| **Rachel Metzger** | | |
| **Justin Gore** | Northwick Park Hospital | Health Services Research and initial Project manager |
| **Principal Research Team** | | |
| **Keith Sephton** | Northwick Park Hospital | UKROC Database development and database manager |
| **Heather Williams** | Northwick Park Hospital | Senior data analyst |
| **Alan Bill** | Northwick Park Hospital | NHS Accountant and analyst |
| **Diana Jackson** | King's College London | Coordination of report preparation |
| **Paul Bassett** | Stats Consultancy | Statistical support |

| | | |
|---|---|---|
| **Patient and Public Involvement (PPI)** | | |
| **Ann and Steve Harris** | Chairs of the PPI Group | |
| **The Buckley Family** | PPI Group | |
| **The Macaffey Family** | PPI Group | |
| **Others** | | |
| **Norman Keen** | Clinical Reference Group for Specialist Rehabilitation | |
| **Jane Pye** | Clinical Reference Group for Specialist Rehabilitation | |
| **Department of Health, NHS England, HSCIC** | | |
| **Peter Howitt** | Department of Heath Payment by results team | Tariff Development |
| **Adebayo Adekalyoaja,** | HSCIC | Casemix development |
| **Chris Foster McBride** | HSCIC | Casemix development |
| **Paula Monteith** | HSCIC | Casemix development |
| **Austrasalian collaborators** | | |
| **Frances Simmonds** | Australasian Rehabilitation Outcomes Centre | Expertise in national dataset development and management |
| **Tara Stevermuer** | Australasian Rehabilitation Outcomes Centre | Statistical analysis of AROC dataset |
| **Prof Barbara Singer** | University of Western Australia | Facilitation of data collection in Perth |
| **Prof Fary Khan** | Royal Melbourne Hospital | Facilitation of data collection in Melbourne |
| **Prof Natasha Lannin** | La Trobe University | Facilitation of data collection in Melbourne |
| **Kristylee Sharp** | Brightwater Group | Data collection in Perth |
| **Karla Seaman** | Brightwater Group | Data collection in Perth |
| **Prof Kath McPherson** | Auckland University of Technology | Facilitation of data collection in New Zealand |

# Table of Contents

# List of Tables

# List of Figures

# List of Appendices

# List of Abbreviations

| | |
|---|---|
| **ADL** | Activities of Daily Living |
| **AHP** | Allied Health Professional |
| **AN-SNAP** | Australian National Sub-Acute and Non-Acute Patient classification |
| **AROC** | Australasian Rehabilitation Outcomes Centre |
| **BI** | Barthel Index |
| **BSRM** | British Society of Rehabilitation Medicine |
| **CCG** | Clinical Care Group |
| **CfH** | Connecting for Health |
| **CLRN** | Comprehensive Local Research Network |
| **CMG** | Case-mix Groups |
| **CMS** | Case-mix System |
| **CRAFT** | Case-mix Rehabilitation and Funding Tree |
| **CRG-SR** | Clinical Reference Group for Specialist Rehabilitation |
| **CSRS** | Complex Specialist Rehabilitation Services |
| **DeNDRON** | Dementia and Neurodegenerative Diseases Research Network |
| **DH** | Department of Health |
| **DRG** | Diagnosis Related Group |
| **DSRS** | District Specialist Rehabilitation Services |
| **ERP** | Expert Reference Panel |
| **EWG** | Expert Working Group |
| **FIM** | Functional Independence Measure |
| **FAM** | Functional Assessment Measure |
| **FRG** | Function Related Groups |
| **GAS** | Goal Attainment Scaling |
| **HES** | Hospital Episode Statistics |
| **HQIP** | Health Quality Improvement Partnership |
| **HRG** | Healthcare Resource Group |
| **HSR** | Health Services Research |

| | |
|---|---|
| **ICD-10** | International Classification of Diseases version 10 |
| **ICF** | International Classification of Functioning Disability and Health |
| **IC-HSC** | Information Centre for Health and Social Care |
| **InterRAI** | International Research Assessment Instrumentation |
| **IQR** | Inter Quartile Range |
| **IRF** | Inpatient Rehabilitation Facility |
| **KCL** | King's College London |
| **LGRS** | Local General Rehabilitation Services |
| **LNSCG** | London Neurorehabilitation Specialised Commissioning Group |
| **LOS** | Length of Stay |
| **LTNC** | Long term Neurological Condition |
| **MDT** | Multi Disciplinary Team |
| **MS** | Multiple sclerosis |
| **NCA** | National Clinical Audit |
| **NHFS** | National Health Financing Scheme |
| **NHS** | National Health Service |
| **NHS-IC** | NHS Information Centre |
| **NIHR** | National Institute of Health Research |
| **NIS** | Neurological Impairment Scale |
| **NIS-Trauma** | NIS for Trauma |
| **NPDS** | Northwick Park Dependency Scale |
| **NPCNA** | Northwick Park Care Needs Assessment |
| **NPH** | Northwick Park Hospital |
| **NPRNA** | Northwick Park Rehabilitation Nursing Assessment |
| **NPCS** | Needs and Provision Complexity Scale |
| **NPTDA** | Northwick Park Therapy Dependency Assessment |
| **NRES** | National Research Ethics Service |
| **NSF for LTC** | National Service Framework for Long Term Conditions |
| **NWLHT** | North West London Hospitals Trust |
| **NW Thames** | North West Thames health region |
| **OBD** | Occupied bed-day |

| | |
|---|---|
| **OCE** | Oxford Centre for Enablement |
| **OPCS** | Office of Population Censuses and Surveys |
| **OPCS-4** | OPCS Classification of Interventions and Procedures version 4 |
| **PAI** | Patient Assessment Instrument |
| **PbR** | Payment by Results |
| **PCSI** | Patient Classification Systems International |
| **PCT** | Primary Care Trust |
| **PPI** | Patient and Public Involvement |
| **PPIPCO** | PPI and Participation Committee |
| **RCP** | Royal College of Physicians |
| **RCS** | Rehabilitation Complexity Scale |
| **RCS-E** | Extended Rehabilitation Complexity Scale |
| **RCS-ET** | Extended Rehabilitation Complexity Scale for Trauma |
| **RCS-SNH** | Rehabilitation Complexity Scale for Specialist Nursing Homes |
| **R&D** | Research and Development |
| **RRU** | Regional Rehabilitation Unit |
| **RUG** | Resource Utilization Group |
| **SCG** | Specialised Commissioning Group |
| **SCI** | Spinal Cord Injury |
| **SHA** | Strategic Health Authority |
| **SNH** | Specialist Nursing Home |
| **SPSS** | Statistical Package for the Social Sciences |
| **SSNDS** | Specialised Services National Definition Set |
| **SUS** | Secondary Users Service |
| **TARN** | Trauma Audit and Research Network |
| **UKCRC** | UK Clinical Research Collaboration |
| **UKROC** | United Kingdom Rehabilitation Outcomes Collaborative |

# Scientific summary

## Background

Long term neurological conditions (LTNCs) give rise to a complex interaction of medical, physical, cognitive, communicative and psychosocial problems, which present a substantial burden of care to both family members and the services that support them.

- The National Service Framework for LTNCs identified the need for specialist neurorehabilitation services for individuals with complex needs but, at the onset of this programme, there was no agreed definition of 'complex needs' - or indeed what constitutes a 'specialist service'.
- The  Department of Health's clinical data systems did not include rehabilitation and there was no common system for standardising information information on patient needs, costs or outcomes, that could be used to compare different services and identify models that offer best quality and value for money.

## Aims and research questions

Learning from established international models, we set out to provide the evidence to underpin the development of case-mix, accurate patient-level costing and funding models to inform tariff costs under the Department of Health's 'Payment by Results' Scheme.

Building on our previous applied health services research programmes, we aimed firstly, to develop the tools and data to determine the diverse rehabilitation needs of patients with LTNCs; and secondly, to inform the development of a nationally co-ordinated approach to needs-led commissioning and provision of specialist neurorehabilitation services.

### *Key research questions were:*

1. How can we measure individual rehabilitation needs and caseload complexity, and determine which patients need specialist services?
2. How are these needs currently provided for in the NHS and what do they cost?
3. How do we balance resources with outcome to optimise cost-efficiency, and which service models offer best quality and value for money in different neurorehabilitation settings?

## Workstreams and outputs from the programme

The programme had five interconnected work streams that ran concurrently.  Each one had specific objectives and deliverables as set out in this report and detailed in chapters 1-3, but are summarised briefly below.

## *Workstream 1 – Measuring rehabilitation needs and complexity*

**Purpose:** First of all, we required a set psychometrically robust tools to measure rehabilitation needs, inputs and outcomes that were fit for incorporation into a national clinical database, and were suitably adapted for the various settings in which they would be used.

**Methods:** Prior to the start of the programme, the rudimentary tools existed, but information was lacking about their validity, psychometric properties and utility for routine application in clinical settings. We used the Medical Outcomes Standards Framework to evaluate the instruments, to investigate their measurement properties and, where necessary, to adapt them. We explored different options for the evaluation of cost-efficiency.

**Outputs:** At the end of the programme, we have a fully validated set of tools to measure the following parameters at patient level. These are:

- Needs for rehabilitation - The Rehabilitation Complexity Scale and Patient Categorisation Tool
- Inputs provided to meet those needs - The Northwick Park nursing and therapy Dependency Scales, a Medical Activities Assessment
- Outcomes in terms of gains in independence - The UK Functional Assessment Measure and reduction of care needs (The Northwick Park Care Needs Assessment)
- Cost efficiency – the time taken to offset the initial costs of rehabilitation by reduction in the cost of ongoing care needs, as estimated by the Northwick Park Care Needs Assessment.

## *Workstream 2 – Current provision of rehabilitation resources*

**Purpose:** It was then necessary to understand the services that were currently available to cater for patients at different levels of complexity, and the differential rehabilitation resource implications of providing for those needs.

**Methods**: We applied the tools developed in Workstream 1 in a variety of neurorehabilitation settings (including Level1 (tertiary) and Level 2 (Local) specialist rehabilitation services. We assessed caseload complexity across these different settings, and analysed the rehabilitation resources (ie inputs from medical, nursing and therapy staff) that were used by patients at different levels of complexity. We also examined the inputs from medical /surgical specialties other than rehabilitation medicine in hyper-acute rehabilitation settings.

**Outputs:** At the end of the programme, we have a clear picture of where these specialist rehabilitation services are currently provided in England, and the additional resource requirements for managing patients with more complex needs. This information was used to develop the weighted payment model in Workstream 4.

## *Workstream 3 – Learning from international costing and casemix methods*

**Purpose:** In order to develop an appropriate casemix and costing methodology for use in the UK, we looked first the casemix models that have been developed in other countries to determine their applicability for use in the UK.

**Methods**: We conducted a review of the casemix and payment models that have been established in the US and Australia to see if any of them were directly transferable to the UK. Australia's health system provided the closest model to the UK National Health Service. Throughout the programme we worked in close collaboration with the Australasian Rehabilitation Outcomes Centre (AROC) to learn from their experience of developing a national database and casemix system for rehabilitation over the last decade.

Once the UK dataset was sufficiently established, we conducted a comparative case-mix adjusted analysis on data collected in Australia and the UK, using the data elements that were common to both datasets.

**Outputs:** By the end of the programme, we had a clear understanding of the differences between the two countries, both in the populations served and the way that rehabilitation services are provided. These differences confounded direct comparison, and our findings demonstrated that a casemix classification based on the Functional Independence Measure (which underpins casemix in Australia and the US) was not fit for purpose within the UK health system. This confirmed the need to take a different approach to development of a casemix and costing model for specialist rehabilitation the UK.

## *Workstream 4 – Costing rehabilitation programmes in the UK*

**Purpose:** The purpose of this part of the programme was to obtain accurate costing of specialist rehabilitation services in England, and to determine the differential costs of treating patients with different levels of complexity. The aim was to establish a costing and payment model that is fair both to providers and commissioners, that would take account of the increased costs of managing complex patients (but only while they remain complex) and at the same time reward efficiency.

**Methods**: We developed a pragmatic patient-level costing methodology for specialist rehabilitation services, and applied this in different settings to quantify rehabilitation service costs. Using data from Workstream 3, we developed a weighted costing model that reflected the differential treatment costs associated with different levels of caseload complexity.

**Outputs:** At the end of the programme, we had developed a novel commissioning currency in the form of a multi-level weighted payment model based on serial Rehabilitation Complexity scores. The commissioning currency was mandated for use by NHS England in April 2012. We also developed a set of indicative tariffs adjusted for caseload complexity, and supplied NHSE England with an evaluation of the cost impact of implementing them – both overall and at the level of the individual provider.

## *Workstream 5 – National Database Development*

**Purpose:** The final workstream involved the establishment of a nationwide database for centralised collation and analysis of patient episode data from specialist neurorehabilitation services (Levels 1 & 2) in the UK. In addition to providing the commissioning dataset for NHS England, the purpose of the database was to provide national benchmarking data on quality and outcomes, and a dataset that will, in future, be large enough to interrogate in order to identify the approaches that work best for different groups of patients.

**Methods**: The UK Rehabilitation Outcomes Collaborative (UKROC) dataset was established in 2010. In order to be commissioned as a Level 1 or 2 specialist rehabilitation service, all providers must be registered with UKROC, and reporting the full dataset. Analysis of the data collected up to April 2014 provided comparative data on outcomes and cost efficiency across the different service models, and explored the predictors of efficiency and cost-efficiency.

**Outputs:** Since April 2012, the UKROC database it has provided the commissioning dataset NHS England. Only activity submitted to UKROC is counted for reimbursement. UKROC provides sign-posting information to NHS England to support the designation of services into the different service levels described within the Department of Health's Specialised Services definition set.

By the end of the programme, UKROC now collates data from all designated Level 1 (n=15) and Level 2 (n=48) specialist rehabilitation services in England. It also has more limited data from other services such as slow-stream rehabilitation and specialist nursing homes. A total of over 22,000 case episodes has been recorded, and the dataset is now growing at a rate of nearly 5000 cases per year. Quality benchmarking reports are now provided for all services at quarterly intervals.

Our multicentre analysis of 4 year's data confirmed the cost-efficiency of rehabilitation for patients with complex needs. The initial costs of rehabilitation were effectively recouped by savings in on-going costs of care within 19 months of discharge. Patients who were highly dependent on admission, were the most cost-efficient to treat, recouping the costs of rehabilitation in just 13.6 months.

## *Conclusions and impact*

This programme represents a substantial body of research, which has improved our understanding of the rehabilitation needs of patients with complex disability, the resources that are required to manage them, and the outcomes that may be expected. It has also provided the Department of Health with critical information about the costs of rehabilitation services, currencies to provide fair reimbursement for cost-efficient intervention, and the scale of cost savings that may be derived from timely rehabilitation interventions.

This programme has evolved though a time of great change in the NHS. The Health and Social Care Act 2012 introduced the most radical re-organisation and restructuring of the commissioning / funding in the entire history of the NHS. This has provided both opportunities and challenges, as described in Chapter 2.

Throughout the programme we have shared our results through peer-reviewed publications, and presentations to reach a wide audience of stakeholders. The findings and developments produced in the course of this programme have been integrated into the commissioning strategy for specialised rehabilitation services as this has progressed over the seven-year life-time of this programme. They have had major impact on national policy in this area.

Data provided by this programme on service configuration has been used by the British Society of Rehabilitation Medicine (BSRM) to drawn up its standards for neurorehabilitation services. These standards have in turned been taken up and used in the NHS-England Service specification for the designation of services. Our data has supported the development of tariffs and commissioning for specialist rehabilitation.

Early findings from this work highlighted deficiencies in the provision of specialist rehabilitation. There was insufficient service capacity to meet demand and many services had inadequate staffing and resources to manage their caseload. The demonstration that rehabilitation was highly cost-efficient led to an increase in commissioning and, throughout the course of the programme, we have seen a gradual increase in service provision, with corresponding increase in the complexity of caseload managed.

From the end of this programme, funding for the UKROC database has been included in the NHSE commissioning portfolio for 2015/16 and going forward. This contracting arrangement confirms the value that NHSE England places on the outputs of this programme grant for the purposes of commissioning and national benchmarking.

Following a successful new topic proposal to HQIP in 2011, a National Clinical Audit has been developed to evaluate specialist rehabilitation following Trauma. The project will link the UKROC and TARN (Trauma Audit and Research Network) Databases to support tracking of patients as they move from the Major Trauma centres to the Specialist Level 1 and 2 Rehabilitation services.

The programme was centred on specialist neurorehabilitation services, but the approach is relevant for wider application. In the course of this programme we have worked with groups in other areas of healthcare, including the Expert Working Panels involved in casemix and tariff development for palliative care and complex neurological disability in Children. These collaborations have led to two further successful applications for NIHR-funded programmes in those fields.

## *Recommendations for future research*

Although the programme has delivered its key targets, this is still a time of major change and development within the NHS. There is still much to be done, working in continued collaboration with NHSE, Monitor and the HSCIC:

Key recommendations for further research and development include:

1. Further development of datasets, tariffs and commissioning currencies for community-based services, including:
    a. Slow-stream rehabilitation in specialist nursing homes and neuro-rehabilitation services. Although this programme touched on these areas there is still uncertainty about the optimum resources required to manage such patients, and the most appropriate tools for outcome evaluation.
    b. Long-term care and support for patients with complex disabilities, both in home-based settings and institutional care (specialist nursing homes).
    c. Specialist community rehabilitation including home-based programmes, day-centre and outreach services.

2. Development of a national clinical registry for patient with prolonged disorders of consciousness to identify patients, monitor progress and interventions, and to record outcomes, including emergence into consciousness and long term prognosis. The UKROC dataset provides the obvious repository for such information, but will require further development to accommodate this information.

These and other developments will be the subject of a follow-on grant application to continue this important and highly productive applied programme of health services research and development.

# Plain English summary


- Following illness or injury, the majority of patients will make a good recovery, but a small number will be left with complex disability, requiring treatment in specialist rehabilitation services.
- Patients with complex needs are more expensive to treat, but there is evidence that rehabilitation can provide value for money by helping them to regain independence and so reducing the costs of long-term care.
- At the outset of this programme the Department of Health had no systematic way of recording information about rehabilitation services. It did not know where the services were, how many patients were treated, or how much the rehabilitation programmes cost.
- This programme established a national clinical database to collect information on the rehabilitation needs of patients with complex disability, the types of rehabilitation they receive, and the outcomes in terms of improved independence and cost-efficiency.
- The UK Rehabilitation Outcomes Collaborative (UKROC) database now collates data from all Level 1 and 2 in-patient specialist rehabilitation in England, and routinely provides comparative information on service quality and outcomes. The data confirm that rehabilitation is highly cost-efficient, effectively paying for itself within 19 months of discharge.
- We also established a system for accurately identifying the cost of rehabilitation, and paying for it in a fair manner that takes account of the higher costs of managing complex patients but and at the same time rewards efficient practice. This is now used by NHS England to pay for in-patient specialist rehabilitation.
- Future research will focus on equivalent developments for community-based rehabilitation services.

# Foreword

This report represents an immense achievement: that of gathering together a wide range of information on neuro-rehabilitation services, sifting through and assessing that information and then making it available in a form that can be used by professionals in this field to make cost and service provision recommendations and decisions that will profoundly affect the future of people with long term neurological conditions.

In the early days of doing the groundwork, the information gathering process alone was just plain hard work; then, making sense of the data collected – and the wide range of assumptions – was daunting. Finally, achieving a balance between the assessment of needs, costs and outcomes was extremely challenging. This was not for the faint hearted and we stand in admiration of the team that produced this report.

It may not contain the answers to all the issues surrounding neuro-rehabilitation, but it does provide an important set of tools which allows needs, costs and outcomes to be objectively assessed, in addition to demonstrating the cost-savings that timely rehabilitation actions can offer. It also provides the means for professionals to signpost to people like ourselves, the parents of a severely brain injured teenager, the facilities that may provide the best rehabilitation outcomes.

When, twenty years ago, we suddenly found ourselves catapulted into a world of brain injury and rehabilitation, we were lost, distressed and completely adrift. We could find little or no guidance on what to expect for our son, what our next steps should be, what to do for the best or who to turn to. We were extremely fortunate, eventually, to have the support of the team at Northwick Park RRU, but that was pure happenchance. What this report provides, for all professionals working with service users such as our son Robin, is a definitive source of factually based guidance on the resources available, backed by an equitable charging and payment system.

We are extremely proud of our association with the report and very much look forward to seeing this important work extended into the area of community services.

**Ann and Steve Harris**

# Robin's Story

Twenty years ago our son, Robin Harris, was brain injured in an accident. During his long – and still on-going recovery, we have gained extensive first-hand knowledge of the services provided to patients with Acquired Brain Injury. We were therefore pleased to be involved in the development of this report, particularly since specialist rehabilitation helped make what we consider to be *the* most significant contribution to Robin's on-going recovery.

From a personal perspective, the specialist rehabilitation unit was one of the few oases of excellence in the provision of post-brain injury services experienced by Robin that were otherwise too often piecemeal, uneven in quality and sometimes difficult to access.

Back in June 1995 Robin was an active, robust and intelligent 14-year old. Playing with friends in the park he fell nearly 40 feet – *approximately the height of two double-decker buses* – from a tree. He was unconscious. His friends put him in the recovery position and called an ambulance – the paramedics in turn called out the Air Ambulance and Robin was airlifted to the Trauma Unit at the Royal London Hospital, within the so-called 'Golden Hour'.

Robin had suffered a severe, diffuse brain injury – his GCS was 3 – but, fortunately, no fractures or ruptures. Because it was a closed-head injury, a bolt was inserted into his skull to measure and relieve inter-cranial pressure. Later a tracheotomy was performed to help him breathe on a ventilator. Robin was in a deep coma and would remain so for the next two months. Doctors predicted a very bleak future for him and, by implication, for us.

In the Trauma Unit Robin had the highest possible standard of nursing and quality of care – and this undoubtedly saved his life. Unfortunately pressure on beds was severe and, after two weeks, despite still being in a coma and still needing a high level of care, he was transferred to a general hospital children's ward.

In this general ward environment he did not receive the care appropriate to his condition:

- he was given no physical or mental stimulation,
- due to his injury, his body temperature soared – no one noticed until we visited, then he was packed with ice,
- his nasal feeding tube ran dry – no one saw this until a visitor called attention to it,
- his tracheotomy became infected so badly that he required surgery.

After the operation Robin was placed once more on a general ward, again suffering from a lack of appropriate nursing and facilities for brain-injured patients:

- he was left in bed for long periods, with no stimulation,
- his limbs were splinted to prevent contractions, but…
- no attempts were made to begin any physical rehabilitation, even just to be stood up – no-one had the time or the necessary expertise.

Whilst there Robin lost over half his body weight and become very weak and emaciated.

Also, a particular concern of ours at this time was the difficulty in getting clear information about a suitable, logical treatment path which would start Robin on the road to recovery. We eventually discovered the specialist rehabilitation unit at Northwick Park through our own researches – it probably helped that we lived locally. An initial assessment was carried out, but Robin was neither alert nor responsive enough to be considered suitable for admission at that time.

We reluctantly agreed on a move to an interim placement, in an attempt to at least enable a start to be made on a basic programme of physical rehabilitation. There a gastrostomy was performed – as a convenience to enable Robin to be fed without a nasal tube – but the operation went badly wrong. Robin's stomach wall was twisted during insertion of the connecting plug, so that his stomach could not process anything. Poor nursing care allowed his stomach to over-fill, causing him to vomit. Lying flat, he ingested this into his lungs, resulting in pneumonia and peritonitis.

For the second time in 3 months Robin was an emergency admission to an ICU, and we were told he might not survive. Two weeks of intensive treatment eventually saved him, but we were then faced with a repeat of our original dilemma – there was nowhere for him to go where he would receive the level of care and attention that his condition demanded.

He was transferred back to the general hospital children’s ward, where the basic problem – lack of appropriate nursing care and any facilities for rehabilitation – remained. Things were looking bleak but, fortunately, Robin’s enforced stay in Intensive Care had, amongst other things, made him much more responsive and alert.

As a result, on being re-assessed at the specialist rehabilitation unit in November 1995 Robin was accepted. Everything then changed – he started to benefit properly from a standard of care and nursing appropriate to his condition, with experienced professionals interfacing with top neuro therapists who were firm, unshockable and well-versed in the specific requirements of patients with ABI.

Robin spent eleven months in specialist rehabilitation, during which time he progressed from being minimally responsive to outside stimuli, needing to be hoisted into and out of bed and being fed and having to use a ‘Putney’ wheelchair, with head and leg restraints to being alert, aware of his surroundings, interactive and feeding himself, driving his own electric wheelchair, and taking his first faltering (supported) steps.

Whilst in the RRU, Robin made significant advances, both physically and cognitively, which undoubtedly had a profound and beneficial effect on the prospects for his future quality of life. The problem was, where could he go next?

At the time we had no idea of the unique and on-going rehabilitation requirements of people with Acquired Brain Injury. Fortunately the unit’s staff were able to assist us with information on suitable places, and help us assess the most appropriate options. After a lot of research and a number of visits, we decided on slow-stream rehabilitation/ education at the Queen Elizabeth Foundation’s Banstead Place. This met strong opposition from our local Health Authority (HA), who wanted a more ‘cost-effective’ local area solution. We considered their facility, geared towards maintenance rather than promoting recovery, inappropriate – such a move would also have negated the efforts and achievements of his specialist rehabilitation programme. It took the personal intervention of his consultant to convince the HA that Robin had further rehabilitation potential and needed to be in a residential educational establishment – not a 'home for the disabled' – and, as he was a recovery prospect, would not be a drain on the HA's resources.

Robin went to Banstead Place for two years of continuing therapy and care with some basic education and then on to the Centre for Acquired Brain Injury Rehabilitation Unit at the National Star College in Cheltenham. He spent four years there, making great progress in an environment that specifically catered for young people with similar disabilities. On graduating, he received a certificate for being the student who had made the greatest progress in Personal Development.

Almost nine years after his accident, Robin was finally able to move into the care of the Brain Injury Rehabilitation Trust, initially at Milton Keynes and, two years later, at Ely, where he has lived for the last eleven years, still continuing to make a slow but steady recovery while developing the skills he needs towards being able to live more independently.

With the dedicated help of BIRT staff he enjoys shopping in a supermarket for his meals – using the self-checkout – and preparing snacks. He also uses a laptop computer – with a special keyboard to minimise the perseveration effects of his ataxia – for writing e-mails, accessing Facebook and compiling a monthly PowerPoint diary about his various activities.

Despite ataxia and poor trunk control, which affect his balance, Robin has learned to walk with a rollator and feed and dress himself. Although his speech remains severely limited, his intelligence and comprehension are good. Robin is sociable, able to interact easily with people and make his wishes known.

He enjoys taking part in activities which help him develop his physical abilities – one of the earliest things he was able to accomplish, even before being able to walk, was water-skiing using a sit-on board. He currently attends a local gym and regularly sails a trimaran on Grafham Water. A couple of years ago he undertook a special trip to Holland on the Jubilee Sailing Trust's tall ship, SV Tenacious, during which he was hoisted to the crow's nest as part of his watch…

We are convinced that, had Robin not been lucky enough to benefit from the excellent care and attention he received in the specialist rehabilitation unit, at what was a critical stage of his post-intensive care recuperation, his current and future quality of life would have been severely compromised. The treatment he received there provided a solid foundation for all his subsequent progress.

**Ann & Steve Harris**

# Chapter 1: Overview of the report

This chapter outlines the background, justification, aims and objectives of this programme of research and summarises the five component work streams that were carried out to meet its objectives, along with their outputs.

## Background

Long term neurological conditions (LTNCs) give rise to a complex interaction of medical, physical, cognitive, communicative and psychosocial problems, which cause disability for over 1 million people in the UK,[1] and present a substantial burden of care to both family members and the services that support them. Rehabilitation services help to keep people out of acute hospitals and support them in the community, or in their own homes, whilst optimising autonomy and quality of life.

By taking a systematic patient-centred approach to inform the provision of cost-effective neurorehabilitation services, this programme of research addressed key priorities in the NHS Improvement Plan, the National Service Framework (NSF) for LTNCs,[2] and the UKCRC (UK Clinical Research Collaboration) classification strategy.

The NSF for LTNCs in particular highlighted the need for specialist neurorehabilitation services for individuals with complex needs. However, at the onset of this programme there was no agreed definition of 'complex needs', or indeed what constitutes a 'specialist service'. Although many neurorehabilitation services in the UK routinely collected information on outcomes, there was no agreed common system for standardising information on patient needs, costs or outcomes, which could be used to compare different services and identify models that offer best quality and value for money.

Learning from established international models, we set out to provide the evidence to underpin the development of case-mix, accurate patient-level costing and funding models to inform tariff costs under Payment by Results (PbR).[3]

To reflect the full history of this seven-year programme of research, the main body of this report begins with a background chapter, followed by separate chapters setting out the five interconnected work streams that made up the programme. These ran concurrently and each one had specific objectives. The final chapters bring the outputs from each work stream together and present a synthesis of findings along with their implications for future rehabilitation service development.

## Aims

Building on our previous applied health services research programmes, we aimed firstly, to develop the tools and data to determine the diverse rehabilitation needs of patients with LTNCs; and secondly, to inform the development of a nationally co-ordinated approach to needs-led commissioning and provision of specialist neurorehabilitation services.

# Research questions

1. How can we measure individual rehabilitation needs and caseload complexity, and determine which patients need specialist services?
2. How are these needs currently provided for in the NHS and what do they cost?
3. How do we balance resources with outcome to optimise cost-efficiency, and which service models offer best quality and value for money in different neurorehabilitation settings?

# Objectives of each work stream

1. To further develop, test and validate a set of standardised tools to measure individual rehabilitation needs and interventions across a range of different specialist service models and settings.
2. To define case-mix and the complexity of caseload in different services by applying the tools in a variety of neurorehabilitation settings and identifying the rehabilitation resources (medical, nursing and therapy time) that are currently provided to meet these needs.
3. To compare different international funding models and patient-level costing and case-mix methods for rehabilitation.
4. To develop patient level-costing protocols and apply these in different specialist rehabilitation settings to determine the differential treatment costs associated with different levels of caseload complexity in the UK.
5. To establish a nationwide database for centralised collation and analysis of case-episode data on needs, inputs, costs and person-centred outcomes from specialist neurorehabilitation services in the UK. Prospective data collection will inform tariff costs and provide ongoing benchmarking of quality, as well as evaluation of clinical benefits and cost-effectiveness.

# Collaborating organisations

This research programme was undertaken in collaboration with  key organisations:

- The Health and Social Care Information Centre (HSCIC)
- The British Society of Rehabilitation Medicine (BSRM)
- The Australasian Rehabilitation Outcomes Centre (AROC)

# Outputs

The outputs of this programme to date have included the development of NHS-wide case-mix and costing models for use across the range of specialist neurorehabilitation services.  The programme has provided:

- A valid set of tools to describe rehabilitation needs and a nationally standardised system for evaluation of needs inputs and outcomes.
- Accurate patient-level costing data which have underpinned development of commissioning currencies and tariff costs for specialist rehabilitation servicesunder PbR.
- A national centralised database, which provides practice-based evidence regarding high quality and cost-efficient service models to inform future planning of neurorehabilitation services.

Whilst the main focus of this programme has been on specialist neurorehabilitation, the principles and methodology could be extended to other areas of rehabilitation, and to other areas of healthcare e.g. specialist palliative care, where alternative commissioning currencies to fixed case episodes are similarly required.

# Chapter 2: Background

## Setting the scene

Historically, the central focus for NHS research and development has been on acute and emergency services, but there is now a growing agenda to address health and social services support for people with longer term needs. The NSF for Long Term Neurological Conditions (LTNCs)[2] published in 2005 emphasised the need for life-long person-centred care and for rehabilitation services, both in hospital and in the community. Until recently, however, there has been little investment in research to understand the rehabilitation needs of this group of patients or cost-effective services to provide for them.

People with LTNCs have a diverse mixture of medical, physical, cognitive, communicative and psychosocial problems. They have widely diverse goals for rehabilitation, requiring input from different disciplines and different types of service. For example:

- Following sudden onset of neurological illness or injury, patients may require post-acute intensive in-patient rehabilitation to enhance their recovery and maximise their return to independence.
- People with very severe long-term disability may require slow-stream rehabilitation in community-based settings such as specialist nursing homes.
- Individuals with severe cognitive deficits and challenging behaviour may require rehabilitation in neuro behavioural settings.

Accordingly, people with LTNCs have widely differing needs for services, against which the adequacy of service provision must be judged. Unfortunately, these needs are poorly recorded and largely unreported through current NHS information systems, so the epidemiology of 'need', as opposed to 'disease', is not fully understood.[4] This problem is not unique to the UK but represents a challenge for clinical outcomes research across the world.

### *What is rehabilitation?*

According to the Royal College of Physicians report: Medical rehabilitation in 2011 and beyond,[5] Rehabilitation is *"an active, time-limited collaboration of a person with disabilities and professionals, along with other relevant people, to produce sustained reductions in the impact of disease and disability on daily life. Interventions focus on the individual, on the physical or social environment, or a combination of these."*

There is now a substantial body of evidence for the effectiveness of rehabilitation, both in the form of Cochrane reviews[6, 7] and from the non-trial-based literature[8] - and also for its cost-effectiveness.[9]

### *Levels of specialisation for rehabilitation services*

Local general rehabilitation services are able to meet the needs of many, but some people have more complex needs requiring more specialist neurorehabilitation services. It is important to be able to define rehabilitation needs in order to match services to the needs of the individual. On the other hand, specialist programmes are more costly to provide and careful monitoring of outcome is required to demonstrate value for money.[10, 11]

Following recommendations from a report by the Royal College of Physicians in the mid 1980s,[12] rehabilitation services have developed on a three-tiered model of local, district and regional services. At the start of this programme, these three levels of service had not been formally defined, and it was largely left up to the provider to classify themselves. In 2009, however, following the Warner Report,[13] the UK Department of Health for England published an updated National Definition Set [14] (3rd edition), which laid down specific criteria to support formal designation of 'Specialised Services' within in 36 healthcare domains. Definition no. 7 'Complex specialised rehabilitation for brain injury and complex disability'[15] identified three levels of rehabilitation service and four categories of patient needs which are illustrated in *Figure 1.*

- **Level 1 services** – were discrete tertiary specialist rehabilitation services. These are high cost/low volume services taking patients with highly complex (Category A) needs and serving a regional catchment of 1-3 million population.
- **Level 2 services** – were discrete specialist rehabilitation services, led/supported by a consultant specialist in rehabilitation medicine. They take patients with moderately complex (Category B) needs and operate within a more local, district-level catchment population (typically 350–500K).
- **Level 3 services** – were local non-specialist rehabilitation services, for patients with category C or D needs. They were often provided in the context of acute or intermediate care, as opposed to a discrete rehabilitation unit and may be led by consultants in other specialties or by allied health professionals.

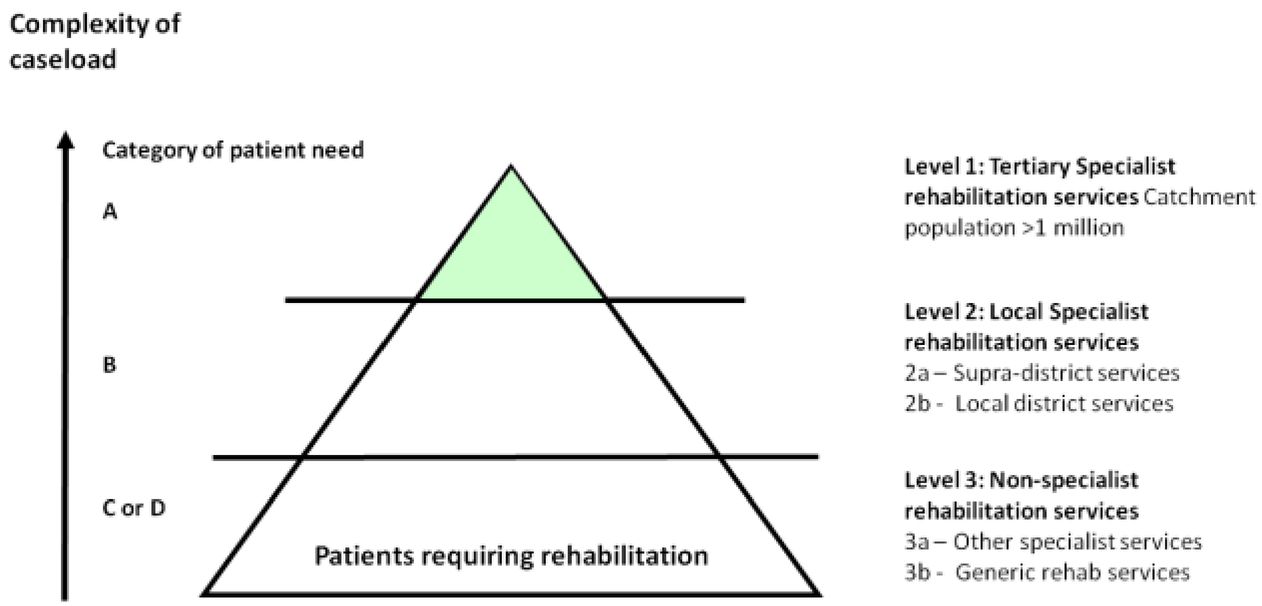



**Figure 1: Levels of rehabilitation services defined by the DH SSNDS third edition**

In general, Level 1 services have been shown to carry a greater proportion of complex cases.[16] However, due to the scarcity of such services, Level 2 services in many areas have evolved to serve a supra-district population (e.g. 600 000–1 million) and also carry a relatively high proportion of complex cases.

In 2010, the British Society of Rehabilitation Medicine published a document entitled "Specialist neuro-rehabilitation services: providing for patients with complex rehabilitation needs"[17] in which these supra-district services were defined as Level 2a services. The revised levels are shown in *Figure 2*. These definitions have continued to evolve throughout this programme. Please see the BSRM website for the latest version of this document (updated 2015)[18] (summarised in Appendix 2.1).

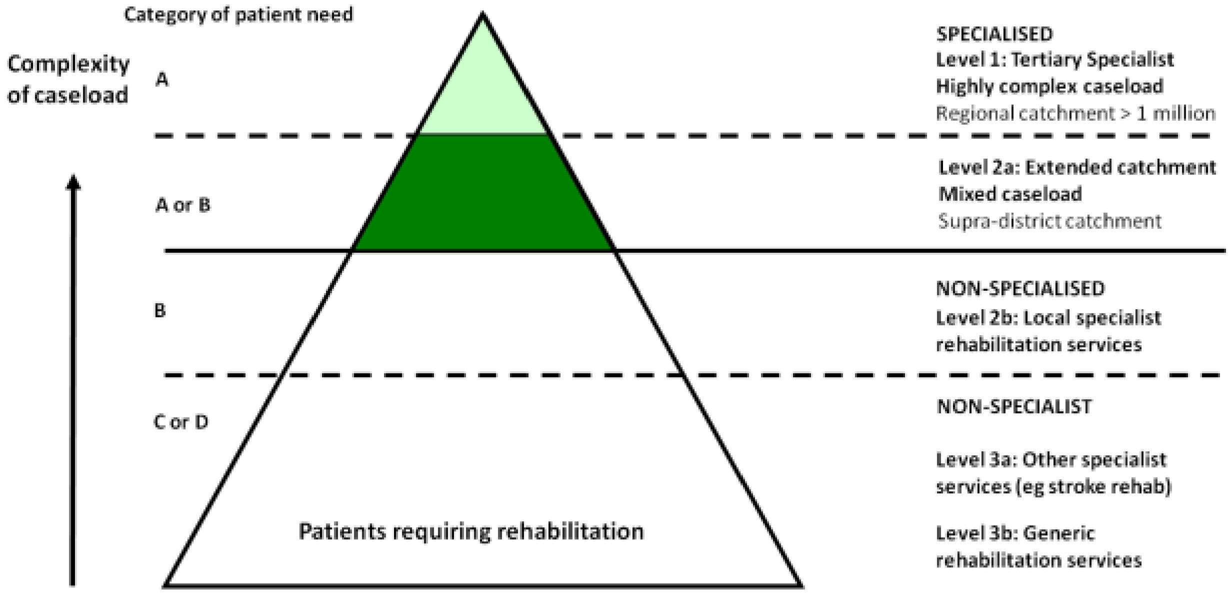



**Figure 2: Further definition of service levels according to the BSRM 2010**

### A note on terminology

We recognise that the terminology 'specialised' and 'specialist' has become somewhat confused, as DH usage has varied over the years, but the terms have sometimes been used loosely and inter-changeably.

- Technically 'specialist' is the umbrella term describing a service that focuses on one particular area of practice.
- In commissioning documents, the DH has sometimes used 'specialised' to describe tertiary services, commissioned on a regional or supra-district basis, whilst 'specialist' may apply to locally-commissioned services.
- In the context of rehabilitation, a 'specialist rehabilitation' service is one led by a consultant trained and accredited in the specialty of Rehabilitation Medicine, as opposed to a 'non-specialist' service that may also provide rehabilitation, but is led by a consultant from another clinical specialty (eg stroke medicine or care of the elderly).

## *Case-mix and payment models – international development*

Health systems around the world are increasingly moving towards payment systems based on a fixed tariff structure for each episode of treatment. The rationale for this approach is to drive up efficiency and to contain costs.[19-21] The challenge, however, is to develop a case-mix classification which adequately accounts for case complexity (i.e. the differential costs of treating patients with different levels of need), and also commissioning currencies that provide equitable risk-sharing between the purchaser and provider.[21]

The term 'Case-mix' is used variously to describe the type of mix of patients treated,[22] in terms of prognosis or resources, depending on the parameter of interest. In the context of healthcare funding models, 'case-mix' systems classify people into groups that are homogeneous in their use of resources, but a good case-mix system also gives meaningful clinical descriptions of these individuals.[23]

Case-mix information has several uses - it can provide the basis for reimbursement, and also for comparing facilities or programmes, practice patterns and patient outcomes. Case-mix information is collected as a standard minimum dataset and, depending on the content of the dataset, it may also be used to benchmark quality and efficiency and for internal management purposes. As well as informing resource allocation, it therefore serves as an information tool that allows policy makers to understand the nature and complexity of healthcare delivery.[22]

Case-mix classification was first pioneered in the United States of America (US) some forty years ago,[24] and Medicare introduced a 'Prospective Payment System' for short-stay acute care hospitals in 1983. The system was based on Diagnosis Related Groups (DRGs), which were designed to measure and classify healthcare activity, based primarily on the diagnosis and the procedures carried out.[25] Each DRG was assigned a 'fixed tariff' payment, as patients within each category were considered to be clinically similar and therefore expected to use the same level of hospital resources. Other countries that have since followed on similar lines to introduce DRG-equivalent systems include Australia, Canada, Nordic Countries, and several European countries including Italy, Germany, Hungary and Denmark.[26]

Since 2002, the NHS has been moving towards a standard 'fixed tariff' payment system along similar lines to the US-style Prospective Payment system. Introduction of the Department of Health's Payment by Results programme[27] in England represented the biggest change in financial flows in the history of the NHS, and has introduced some particular challenges as payment systems change from a 'block contract' model (in which the provider is paid a flat annual fee to provide an agreed level of service) to activity-based or 'cost per case' model (which reimburse providers on a case-by-case basis). Payment is made according to a standard national 'tariff' – the price tag attached to any unit of care, wherever it is delivered.

A trial application of the US-DRGs in the UK acute care sector demonstrated that they did not accurately fit the case-mix of the NHS.[28] Instead, the UK scheme introduced a national episode-based tariff for healthcare treatments, based on a case-mix classification of Healthcare Resource Groups (HRGs), which are derived from a combination of diagnosis (ICD-10) and procedure (OPCS) codes [29, 30] (see glossary of terms for further explanation). Like DRGs, HRGs are groups of conditions and interventions that are intended to have similar resource implications. The latest iteration (HRGs version 4) has been expanded significantly to underpin standard tariff payments for NHS treatments.[31] Tariffs for each HRG are normally determined through analysis of 'reference costs' - the average cost of treatment spells (episodes) within each category, reported by healthcare providers.[27]

As in other episode-based funding models, the system was intended to reward additional activity and give an incentive to reduce length of stay, so driving up cost-efficiency. However, many specialties have recognised the need for increased 'granularity' of the classification to account for the additional costs of treating people with more complex needs. The Department of Health for England also accepted the evidence that shorter lengths of stay do not always equate with efficiency,[11, 32] and so started to explore alternative payment models for management of long term conditions, for example in the areas of mental health, palliative care and rehabilitation. The introduction of mandatory tariffs was therefore put back in those areas, pending further refinement of the case-mix model.

## Case-mix in rehabilitation – international models

When the US case-mix system was first introduced in the 1980s, medical rehabilitation was excluded from the US-DRGs because it was recognised that rehabilitation in-patients could not be classified reliably by diagnosis alone.[33] The level of functional dependency was considered to be a better cost-indicator, and in the 1990s a classification system based on function (as defined by the Functional Independence Measure (FIM)[34]) was developed instead.[35] This system of 'FIM-function-related groups' (FRGs) was subsequently re-derived to predict total rehabilitation costs and re-named 'case-mix groups' (CMGs). These form the current basis for reimbursement by Medicare for in-patient rehabilitation[21] in the US.

Similarly when Australia introduced case-mix funding based on the Australian National Diagnosis-Related Groups (AN-DRGs), it was recognised that the costs of sub acute care (including rehabilitation medicine, palliative care, psycho-geriatric and elderly care evaluation) were not well described by this system.[36] As in the US, function-related groups were identified as a better way to determine the differential cost of rehabilitation episodes.[37] Two separate function-related case-mix systems were developed to classify patient episodes for different levels of reimbursement for rehabilitation.[38 39] The more sophisticated of these two systems is the AN-SNAP (Australian National Sub-Acute and Non-Acute Patient) Classification based on the FIM.[39] This now forms the basis for the activity based funding (ABF) model for sub-acute care services (including rehabilitation) introduced under their National Health Reform Agreement in 2011.

### Case-mix in rehabilitation – the UK

The UK has a very different health culture from that in the Australia or the US. The National Health Service (NHS) provides the most comprehensive publicly-funded healthcare system in the world, with the majority of healthcare being provided free at the point of delivery. In addition, a closely integrated social services system provides life-long care and support, which is free to all individuals who are unable to pay for it themselves. In the light of this on-going responsibility, maximising functional independence is highly valued.

As already noted, the move towards an episode-based single tariff for healthcare treatment (Payment by Results (PbR))[3] is intended to reduce length of stay and drive up cost-efficiency. However, it is unclear how tariffs based on defined episodes of treatment will be applied to life-long care requiring inputs that may be unpredictable in intensity and duration; nor how this one-size-fits-all approach will take account of people with more complex needs. Explored here in the context of neurological rehabilitation, the issues are also relevant across the broader spectrum of chronic disease management in clinical practice.

The implementation of episode-based tariffs for PbR in rehabilitation poses a number of challenges:

- Existing coding systems (e.g. ICD10) cater poorly for rehabilitation, so the Department of Health (DH) had almost no retrospective data on which to estimate activity and base tariff costs.
- Where data were available, reference costs for rehabilitation show very high levels of variance due to the marked diversity of needs for rehabilitation among patients within each diagnostic grouping.

A further challenge is presented by variation in caseload across different levels of service specialisation. Fixed tariffs are based on the assumption that providers carry a similar caseload. As noted above, rehabilitation services in the UK are provided on a three-tiered model of local general (Level 3), district specialist (Level 2) and tertiary specialist Level 1) services. To gain the advantage of critical mass, complex cases have been clustered into the Level 1 centres with specialised staff and facilities.[40] This clustering is feasible in the UK because of the relatively small geographic distances involved. However, it distorts the reference costs, as caseload complexity varies across the different levels of service. This has led to calls for more accurate 'patient level costing',[41] and a case-mix classification based on complexity of needs for rehabilitation.

Following a national consultation, the timescale for introduction of PbR was extended.[42] The revised plans for implementation of PbR focussed on patient-level costing in models of best practice, and identified the need to explore other internationally established and costing models for potential application in the NHS. Rehabilitation was an identified priority for case-mix development, but it was recognised that targeted research would be needed to underpin this development.

## *What we set out to achieve*

Within this programme we set out to develop a nationally consistent approach to case-mix, costing and commissioning of specialist rehabilitation services. We sought to:

- Establish a costing and payment model that is fair both to providers and commissioners, that would take account of the increased costs of managing complex patients (but only while they remain complex) and at the same time reward efficiency.
- Establish the UK Rehabilitation Outcomes Collaborative (UKROC) database as the national clinical database providing benchmarking of quality and outcomes in relation to needs and provision of specialist rehabilitation services in the UK.
- Provide a dataset that will, in future, be large enough to interrogate in order to identify the approaches that work best for different groups of patients.

Firstly, we needed a valid measure of rehabilitation needs. In order to understand (and ultimately cost) the interventions provided in the different levels of service, we require a measure of rehabilitation 'inputs'. We would expect to find that patients with more complex needs receive higher levels of input, at correspondingly greater costs. But to demonstrate that this increased investment in rehabilitation is worthwhile, we need to measure outcomes such as improved independence that would ultimately lead to savings in other service costs.

In any service providing rehabilitation for severely disabled patients, the principal determinant of cost is staff time. This programme has developed and implemented a practical set of tools which can be used to measure a patient's specific needs for care and therapy, and their impact in terms of nursing, therapy and medical staff time. These are designed to help service managers to adjust staffing levels to meet the demands of the case-load, and to provide a basis for comparison of different service models in terms of both effectiveness and cost-efficiency.

If systematically applied and recorded across a range of services, these tools can provide data to assist in service planning. The programme represents the first attempt to set in place centralised gathering of clinical data in this area, with intent to provide generalisable and useable knowledge to clinicians in the field, as well as those working in other areas of long-term illness.

## *Meeting the challenges – the changing scenery of the NHS*

A key challenge to any applied HSR programme is ensuring that the research findings and outputs are fully implemented into continuing NHS practice. With this in mind, this major national programme was set up through a cross-institutional partnership (representing three NHS trusts, three UK universities and one international partner) in close collaboration with a number of key organisations, the NHS Information Centre (as it was known then) and the British Society of Rehabilitation Medicine (BSRM).

The original programme was timed to deliver the research and development work to underpin implementation of PbR in rehabilitation by 2011/12, and we worked in close collaboration with the DH's PbR team throughout. However, we recognised that NHS plans and strategy are apt to change with successive governments, and that policy could change in the lifetime of this grant. We therefore adopted a flexible approach to ensure that it delivered the appropriate outputs to best inform reimbursement for rehabilitation services in the light of prevailing NHS policy and priorities of the time.

In the event, the programme had to evolve through a period of even greater change than we imagined. Following election of the Conservative-Liberal Democrat coalition in 2010, a white paper entitled *"Equity and Excellence: Liberating the NHS"* (DH 2010), the Health and Social Care Bill was introduced to the House of Commons in January 2011 and eventually received Royal Assent on 27.3.2012.

The resulting Health and Social Care Act 2012 introduced the most radical re-organisation and restructuring of commissioning / funding in the entire 63-year history of the NHS.

- It removed from the Secretary of Health responsibility for the health of citizens in England.
- It abolished NHS primary care trusts (PCTs) and Strategic Health Authorities (SHAs) and transferred responsibility for health care funding to Clinical Commissioning Groups (CCGs), led by the general practitioners (GPs).
- NHS England (NHSE) was established as an executive non-departmental public body of the Department of Health, to oversee the budget, planning, delivery and day-to-day operation of the commissioning side of the NHS in England. It also commissions directly the specialised services that were previously commissioned by the Specialised Commissioning Groups (SCGs). In 2012, these services were commissioned nationally through ten of NHS England's 27 Local Area Teams. They accounted for around £11.8 billion of annual spending, or around 12 % of the overall NHS budget.
- Monitor is a further executive non-departmental public body of the Department of Health. Originally established in 2004 as the regulating body for NHS Foundation Trusts, it was assigned further duties by the Health and Social Care Act, which included setting prices for NHS-funded care in England.
- Under the new structure, therefore, the DH PbR team was abolished. NHS England became responsible for the development of currencies and payment models, while Monitor was responsible for setting and development of tariffs.
- The NHS Information Centre changed from a special health authority to become the Health and Social Care Information Centre (HSCIC), a further executive non-departmental public body of the Department of Health. It continued to be England's central, authoritative source of health and social care information and to store and analyse data on all hospital activity in the NHS in England. It also took over the functions of the NHS Connecting for Health programme, which had previously been responsible for IT infrastructure, including the OPCS coding system and delivery of the NHS IT programme.

2012/13 became a transition year during which these changes were put in place. The Specialised Services National Definition Set (SSNDS) was subject to review and overhaul – the 36 services definitions being expanded to a total of 75 in the course of that period. Sixty-five Clinical Reference Groups were established under five programmes of care to bring specialist expertise and advice, together with the views of patients and carers. Although originally conceived as assurance groups, CRGs became the key delivery mechanism for development of specialised services contract products for 2013/14. They were tasked with the overall responsibility of ensuring that services specifications and clinical commissioning policies were delivered on time and to a high quality standard.[43] Under these reforms the SSNDS definition No 7 '*Brain Injury and Complex Rehabilitation*'[15] was renamed to become "*Specialist rehabilitation for patients with highly complex needs*".[44]

## Designation of Level 1 and 2 services

NHSE was tasked with identifying and designating the specialised rehabilitation services to be directly commissioned under this specification. As the DH held no data by which to identify this activity, NHSE relied on the UKROC database for the identification of eligible service providers and activity. As a result, from April 2013 the UKROC database became the NHSE commissioning dataset. As there was not a Level 1 service in every area, the specification included patients with category A needs undergoing rehabilitation in either a Level 1 or Level 2a service as an interim arrangement. It was anticipated that with local commissioning barriers removed, Level 2a services in those areas without a Level 1 service would increase their catchment and complexity profile to an extent where they would be eligible for designation as a Level 1 service within 3 years.

### Opportunities and Challenges for this programme

These sweeping changes have required considerable flexibility as the focus needed to change as the plans for re-structuring were developed, finalised, and implemented. Although the overall direction and final outputs are similar, the ordering has had to vary as different tasks were prioritised by NHSE and Monitor. In addition, some aspects of the programme were developed more quickly than anticipated, but other areas have been deferred.

This has brought both opportunities and challenges for this research and development programme:

- On the positive side, it provided an opportunity to embed principles into policy from the outset within the specialist services. For example, the service specification leaned heavily on the previous SSNDS definition and reflected the quality standards published by the BSRM[17, 45] regarding configuration and staffing of rehabilitation services. Registration with UKROC, and submission of the full UKROC dataset for each admitted patient episode, became a mandated requirement for all specialist (Level 1 and 2) rehabilitation services in England. This meant that national roll-out of data collection occurred sooner than expected, and more in-patient centres were recruited early in the programme.

- On the negative side, it placed a hold on the further development of currencies and tariffs while Monitor and NHS England staff were appointed and worked into their new roles. Even though the CRG and NHS England made a strong recommendation to Local Area Teams to implement the mandated currency and to work towards the indicative national tariffs, steady state commissioning in 2013/14 and 2014/15 effectively meant that commissioners could do as they chose. Some chose to implement the tariffs but others did not. In addition, tariff development for community based rehabilitation was expected to come to the forefront at an earlier stage, but has been deferred by PbR and Monitor to an extent where they could not be progressed within the life time of this programme.

These impacts will be discussed in more detail in subsequent chapters, which will summarise any amendments to the methodology as a result of these changes.

# Chapter 3: Overview of our approach

The programme of research comprised five parts, which ran concurrently.

- Parts 1 and 2 provided the standardised tools for measuring rehabilitation needs in people with LTNC, defining case-mix and the complexity of caseload in different services, and the resources (in terms of medical, nursing and therapy time) that are currently provided to meet these needs.
- Parts 3 and 4 explored different funding models and patient-level costing methods that could be used to determine the treatment costs associated with different caseload complexity.
- Part 5 established a national database to gather systematic prospective case-episode data from neurorehabilitation services. As the tools developed in parts 1-4 stabilised, data were incorporated on needs, inputs, costs and person-centred outcomes, and these have been analysed to provide a UK-wide perspective on the following questions:
    a) What are the costs of providing neurorehabilitation for patients with different levels of rehabilitation need in different settings?
    b) What patient characteristics determine the type of rehabilitation services they require?
    c) Which models of care offer best value for money in different settings?

A sixth part covered dissemination to support implementation through guidance and policy development.

*Figures 3 and 4* illustrate the various packages of work and their timescale across the seven year duration of the research programme.

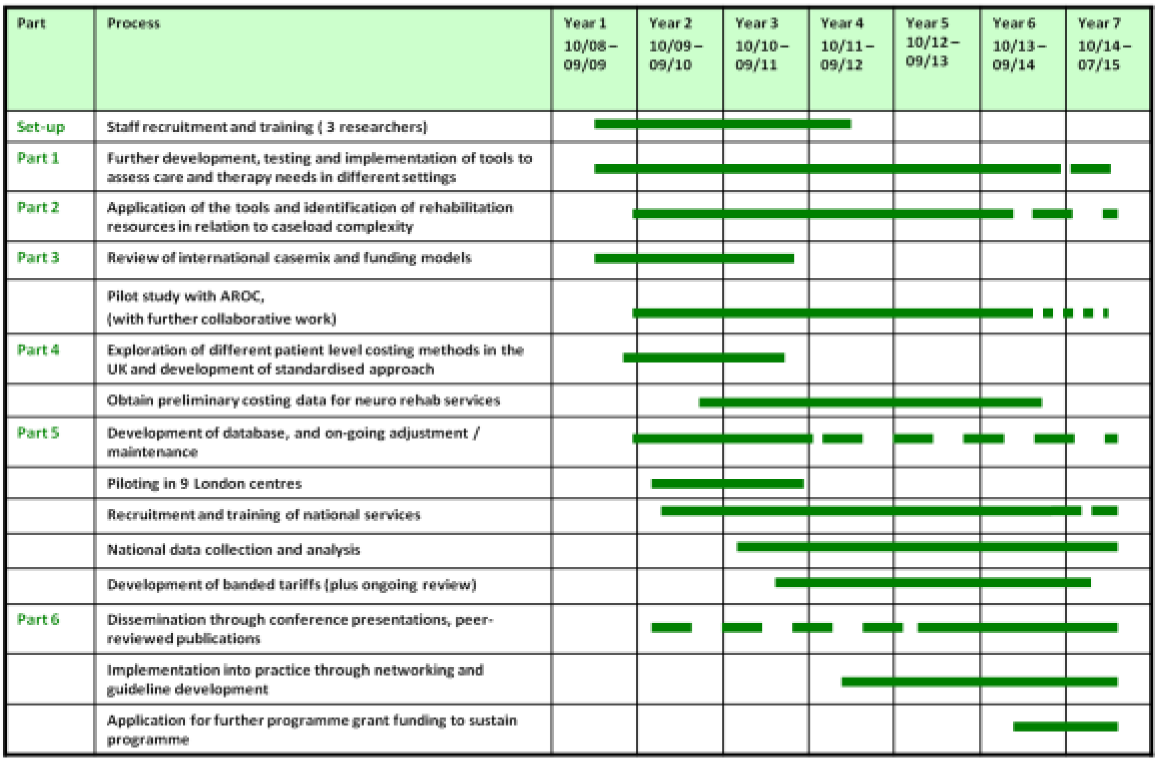



**Figure 3: Gantt chart showing processes across the seven year timescale**

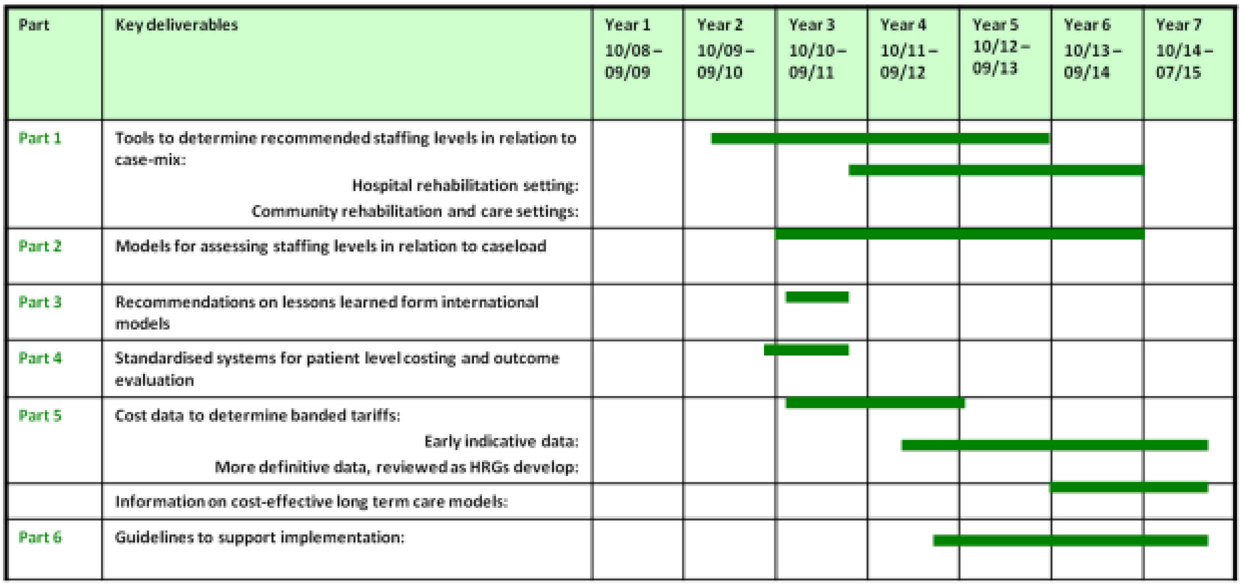



**Figure 4: Gantt chart showing key deliverables across the seven year timescale**

# Key deliverables

| | | |
|---|---|---|
| **Part 1** | **Measuring rehabilitation needs and complexity – tool development** | |
| ***Objective*** | ***To further develop and validate a set of standardised tools to measure individual rehabilitation needs and case-load complexity in neurorehabilitation settings.*** | |
| **Key Deliverables** | ○ A set of psychometrically robust tools to measure needs inputs and outcomes, that are fit for incorporation into the UKROC database. | **Achieved** see Chapter 4 |
| | ○ The tools to be adapted and customised where necessary for the various settings in which they will be used. | **Achieved** see Chapter 4 |
| **Part 2** | **Current provision of rehabilitation resources** | |
| ***Objective*** | ***To apply the tools in a variety of neurorehabilitation settings and to identify the rehabilitation resources currently provided in relation to caseload complexity.*** | |
| **Key Deliverables** | ○ Analysis of staff time in relation to complexity of rehabilitation needs to inform the weighted payment model. | **Achieved** see Chapter 5 |
| | ○ Analysis of caseload complexity in relation to resources (medical, nursing and therapy staff) to inform designation of specialist rehabilitation services by NHS England. | **Achieved** see Chapter 5 |
| | ○ A system for identifying patients with category A needs (requiring specialist (Level1/2a) services. | **Achieved** see Chapter 4 |

| | | |
|---|---|---|
| **Part 3** | **Learning from international costing and case-mix methods** | |
| ***Objective*** | ***To compare different international costing and case-mix methods.*** | |
| **Key Deliverables** | ○ A review of international systems for case-mix and payment models in rehabilitation. | **Achieved** see Chapter 6 |
| | ○ Comparative analysis of UK and Australian datasets across a range of long-term neurological conditions - case-mix (according to the Australian (AN-SNAP) classification, length of stay, outcomes (FIM) and costs will be complete based on the 2013/14 UKROC data. | **Achieved** see Chapter 6 |
| | ○ Development of the national training and accreditation programme for the UK FIM+FAM and other tools in the UKROC dataset to ensure their consistent application. | **Achieved** see Chapter 8 |
| **Part 4** | **Costing rehabilitation programmes in the UK** | |
| ***Objective*** | ***To develop patient level-costing protocols and apply these in different settings to determine the differential treatment costs associated with different levels of complexity.*** | |
| **Key Deliverables** | ○ Development and implementation of a novel weighted bed-day costing and payment model for specialist rehabilitation to derive indicative tariffs for specialist rehabilitation. | **Achieved** see Chapter 7 |
| | ○ Implementation of the model to obtain accurate costing data to underpin a set of tariffs for specialist Level 1 and 2 rehabilitation services. | **Achieved** see Chapter 7 |
| | ○ Evaluation of the cost impact of mandating the tariffs will have been provided to NHS England. | **Achieved** see Chapter 7 |
| **Part 5** | **National database development** | |
| ***Objective*** | ***To establish a nationwide database for centralised collation and analysis of patient episode data from specialist neurorehabilitation services (Levels 1 & 2) in the UK.*** | |
| **Key Deliverables** | ○ All Level 1 and 2a services will be registered and routinely submitting high quality case-episode data to the UKROC dataset. | **Achieved** see Chapter 8 |
| | ○ Routine reporting and feedback systems established to provide contributing centres with activity analysis and benchmarking of outcomes. | **Achieved** see Chapter 8 |
| | ○ Final analysis of cost predictors and outcome in relation to complexity and other factors will have been undertaken on the cleaned 2013/14 dataset. | **Achieved** see Chapter 7 |
| **Part 6** | **Guidance and policy development** | |
| ***Objective*** | ***Dissemination to support implementation through guidance and policy development.*** | |
| **Key Deliverables** | ○ The UKROC database will be fully established as commissioning dataset for Level 1 and 2 specialist rehabilitation services across England, and as the vehicle for administration of the multi-level weighted bed-day tariffs. | **Achieved** see Chapter 9 |
| | ○ Data registry status for the UKROC database will have been applied for. | **On-going** |
| | ○ Linkages between UKROC and TARN as the basis of the National Clinical Audit (NCA) for Trauma and Complex neurorehabilitation will have been defined and the further work needed to provide the NCA will have been scoped and applied for through the HQIP scheme. | HQIP NCA programme contracted only from 1.6.2015 |

# PPI and Stakeholder engagement

## *From the start:*

The conceptual origins of the UKROC dataset collected within this programme are rooted in the NSF for Long Term Neurological Conditions, and its associated LTNC Dataset, developed by the NHS Information Centre. Both of these were developed through Expert Reference Groups with very strong user representation – indeed both panels were chaired by user representatives. The current UKROC dataset is the *'In-patient rehabilitation'* subset of the LTNC Dataset, and tools that are being developed for use in the community are part of its '*Community Rehabilitation'* subset.

## *The dedicated UKROC user group:*

A dedicated UKROC User Group was set up at the start of this programme, chaired by Ann and Steve Harris. The group consists of patients, carers and family members who have used specialist rehabilitation services. Group membership has changed somewhat over the course of the programme as new members have joined. So far as possible, we have attempted to recruit user representatives who have experienced a range of specialist services. For example, Ann and Steve Harris are parents of Robin (now aged 30) who had a catastrophic brain injury at the age of 14 (see the Foreword to this report). Over the last 15 years they have experienced rehabilitation and care in a range of in-patient and residential settings including two Level 1 in-patient specialist services, three specialist slow-stream programmes, and now finally a personalised group home setting. Other members of the group have similarly experienced rehabilitation in both inpatient and community settings.

In its first meeting, the UKROC User Group scoped out the key areas of user involvement in this programme, and also links with other user groups:

- The main areas of interest to users were the evaluation of needs in relation to service provision, which would be addressed in analyses in the final phase of the programme.
- User involvement in the earlier phases focused on development aspects - particularly of tools that would be used in the community-based phases to gather data from patients and carers.
- The technical aspects of service costing and case-mix development were of interest to the users, although there was less direct opportunity to influence this area from within the users group itself.

Contact with members of the group has taken a number of forms throughout the programme.

- Formal user group meetings have been held to update the members, receive their views and share ideas about future directions.
- The chairs have also joined meetings of the Programme Steering Group, and Ann Harris also joined the NHSE Clinical Reference Group for Specialist Rehabilitation (CRG-SR).
- In between, there have been contacts through e-mail, telephone and individual discussions.

## *Other user inputs*

We also looked for opportunities to link in with existing user panels to maximize efficient use of time. For example, the established user/carer advisory panels from our research programme based at King's College London provided very helpful input for community-based tool development, in particular, self-complete versions of tools, including the Northwick Park Dependency Score and the Needs and Provision Complexity Scale. This feedback was gathered both through a range of media including discussion groups and through individual e-mail/telephone contact.

Further involvement of users has occurred as patients have completed self-report questionnaires. For example, follow-up telephone interviews were conducted following receipt of self-report versions of the NPDS. Among the issues raised were presentation of the tool and timeliness to complete, which led to adjustments. Areas of ambiguity were identified through absent or incongruous responses, or through annotations made by the patients / their carers on the questionnaire itself. All of these have helped to develop the final instruments into an accessible toolset which is understandable by the majority of people.

### Engaging the healthcare professionals

Another important sector of the public who will use the outputs of the programme are the clinicians and other health professionals who collect the information on the ground. It was always expected that the UKROC database would continue in perpetuity as the commissioning and national benchmarking dataset after the end of the programme. Clinicians will only engage, however, if they can see the direct benefits and use the data in their day-to-day practice and clinical-decision making. Through our regional workshops and road shows, we have worked closely with staff on the ground to develop the dataset and tools in ways that make them accessible and clinically useful.

Key factors that have ensured a high level of clinical engagement with this programme have been:

1. The direct link between payment for services and data reporting through UKROC has ensured the engagement not only of the clinical teams but their managers, to provide support with data collection and computer entry.
2. Embedding tools within the UKROC software to support clinical use of the data for decision-making and team reflection.
3. Proactive training of consultants to provide strong clinical leadership within their teams to support data collection.

An article published in Disability & Rehabilitation describes some of the steps we have taken to engage the hearts and minds of clinicians in collection of the UKROC dataset.[46]

### Policy development - engagement of commissioners and PPI

Systems for commissioning specialist rehabilitation, including tariff modelling and early database development, were pioneered in the early phases of the programme by the London Neurorehabilitation Specialised Commissioning Group (LNSCG). This group included service users, and again a systematically developed PPI strategy was integral to the work of that group. During the transition to national commissioning, the CRG-SR carried forward the service specification and data requirements that were developed by the LNSCG, which have formed the basis for the programme. We engaged actively with the Local Area Team Commissioners, both through the CRG-SR and through our regional road shows. The UKROC team has played a very pro-active role - providing information and support for implementation of the commissioning model, and seeking feedback on the most useful information to provide in benchmarking reports etc.

The CRG-SR includes services users on its membership, who played a highly interactive role in the developments of the service specification and the on-going strategy for PPI. They also prepared a patient and public-friendly summary of the specification for posting on the CRG-SR website.

### The wider platform

Throughout the programme we have presented at conferences and public meetings at which public and patients invariably form part of the audience and interaction. Importantly, 'PPI' in this context includes not only service users and their families, but people from many different backgrounds (including politicians, journalists and other media workers, economists, teachers, and of course health and social care professionals / commissioners from other areas of clinical practice etc). These individuals have engaged with us over issues raised by those presentations, not only giving us their feedback, but sharing their knowledge and insights, and pointing us in the direction of other tools and resources which we would not have found through our searches of the standard clinical and academic literature.

Within this particular programme, much of the most productive interaction has been overseas – particularly in Australia, New Zealand and the US where development of case-mix and activity-based funding is 10-20 years ahead of developments in the UK. Our correspondents have shared their first-hand experiences of the successes and failures of other systems, which has played a fundamental role in shaping this programme. In this respect we cannot emphasise enough the importance of regular field-based interaction, both at formal conferences and simply interacting with people on the ground. The internet and electronic communication has revolutionised world-wide communication, but nothing can replace face-to-face interaction when teasing out the more challenging data issues. We are fortunate to have been able to take advantage of both approaches during the evolution of this programme.

### Challenges of PPI involvement

Long-term neurological conditions present as a very diverse group of conditions, and patients have widely varying needs and goals for rehabilitation. Naturally, service users and their families become expert in management of their own condition, but may sometimes find it hard to extrapolate that experience to other situations. We have therefore tried to engage with a wide cross section of patients and their families.

One of the most significant challenges is involvement of patients themselves. By definition, individuals with complex disabilities involving physical, cognitive and communicative deficits may have limited capacity to engage directly, and are often excluded from discussions. In this programme, we have taken an inclusive approach and have worked hard to involve patients who have experienced cognitive and communication difficulties, as well as those with purely physical deficits. For example within the UKROC Users Group we now have two patients with dysphasia, and also two who have emerged from vegetative or minimally conscious states.

### Going forward

Coming into the final stage of this programme and beyond, we will continue to work along all of the lines described above to ensure maximal PPI throughout the reporting stage as we finalise the report. We are currently in the process of updating the UKROC pages on our website and plan to include a specific section for PPI. The service users will play a vital role in developing the material. Once the report has been finalised, we plan to produce a user-friendly summary of the main findings and plans for implementation.

# Data collection – The UKROC database

The national dataset for specialist rehabilitation is collated through the UK Rehabilitation Outcomes Collaborative (UKROC) database. The database provides:

- Centralised collation and analysis of patient episode data from all Level 1 and Level 2 specialist in-patient neurorehabilitation services in England.
- Automated calculation of the weighted payment for each episode from serial RCE ratings and length of stay data.
- Case-mix and costing to inform development of national tariffs.
- National benchmarking of services, and prospective cohort analysis to evaluate effectiveness and cost-efficiency in different service models.

The tools by which this information is gathered have been previously developed and validated, but are undergoing further development in the course of the programme to adapt them for the various different contexts in which they are to be applied.

1. **Complexity of rehabilitation needs** is captured by the Rehabilitation Complexity Scale (RCS) - a simple measure of the requirements for care, nursing, therapy and medical management.[47] The RCS has been through several iterations in the course of its development throughout this programme, but the version used for the majority of analysis in this report is the Extended RCS (RCS-E-v12).[48]

2. **Rehabilitation inputs** are captured by the Northwick Park Dependency scales. These are used to identify the rehabilitation resources provided in relation to caseload complexity.
   - The Northwick Park nursing Dependency Scale (NPDS)[49, 50] provides an assessment of care and nursing needs and translates by way of a computerised algorithm to an estimation of nursing and care staff hours.[51]
   - The Northwick Park Therapy Dependency Assessment (NPTDA)[52] is the therapy equivalent, which collates therapy inputs from the multidisciplinary team and also translates by a computerised algorithm into an estimation of therapy hours for each discipline (including medical staff).

These tools are designed to be used either prospectively to capture 'needs' or retrospectively to capture 'inputs' actually provided.[53] In the context of this study, they are applied retrospectively as a measure of the resources (staff time) that were used.

3. **Outcomes:** All units are asked to collect a minimum of standardised outcome data which includes either a Barthel Index,[54] the FIM and/or UK Functional Assessment Measure (UK FIM+FAM).[55] These outcome measures were chosen because previous work has shown that 95% of specialist rehabilitation units in the UK were already using one or more of them as part of routine clinical practice.[56, 57] Recording of Goal attainment scaling (GAS) offers a further non-mandatory option for capturing individualized person-centred outcomes, where providers consider that the standardised measures fail to capture the intended goals for treatment.[58, 59]

In order to minimise the burden of data collection, the UKROC dataset was initially designed to be hierarchical, so that higher volume low cost (Level 2) units could collect just a minimum dataset of five items, whereas the lower volume Level 1 specialist services could reasonably be expected to collect a more detailed dataset. The hierarchical dataset is illustrated in *Figure 5.*

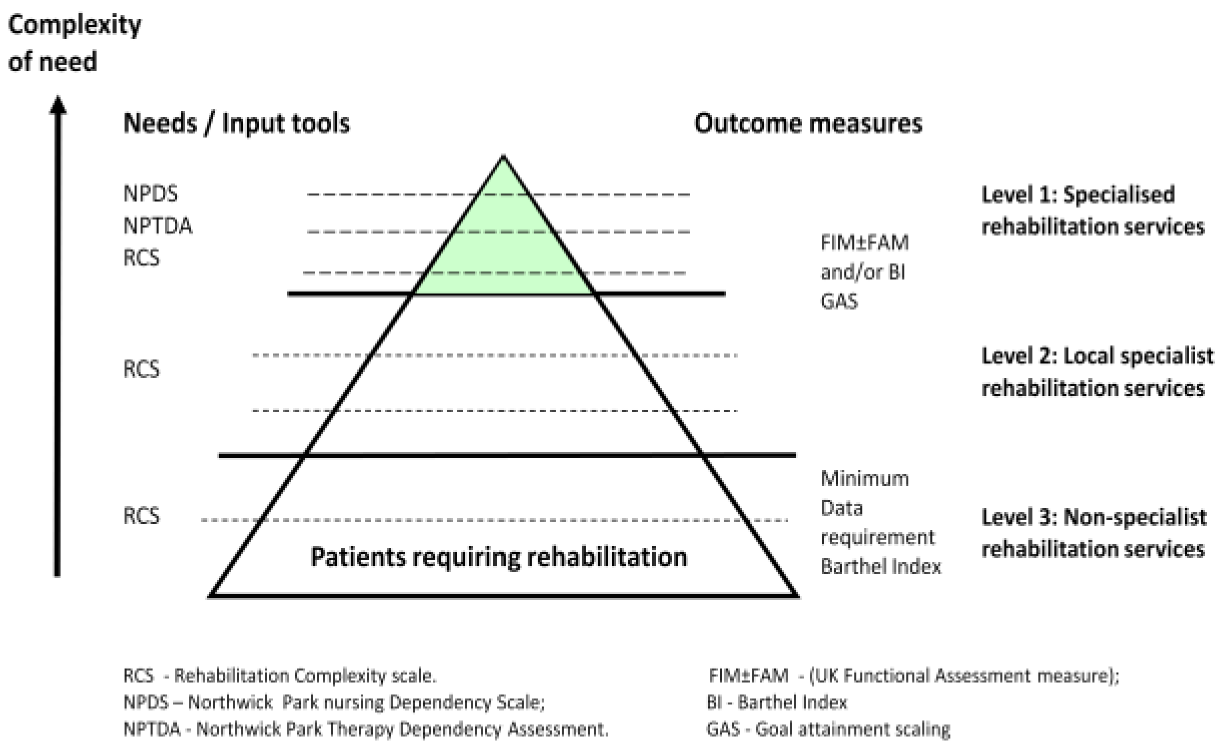



**Figure 5: Tools for measuring complexity of need, inputs and outcomes across the three levels of rehabilitation service provision**

**Legend:** *A hierarchical system of data collection has been developed, so that low cost/high volume services are not over-burdened by data collection. The RCS and Barthel Index offer simple practical measurement of complexity and outcome in the Level 3 and some Level 2 services. For Level 1 services, the Northwick Park Dependency Scales and the Functional Assessment Measure (with optional Goal Attainment Scaling) provide more detailed definition of the needs and interventions as well as opportunities for evaluation of outcomes and cost-efficiency in patients with complex needs.*

In practice, once clinical teams got used to the tools and started to integrate them into their daily practice, the system was found not be unduly onerous. Since NHS England became responsible for commissioning of all Level 1 / 2a specialist rehabilitation services in England, UKROC has become the national commissioning dataset. Only data collected and reported through UKROC is eligible for funding under this programme. Completion of the full UKROC dataset has been a mandated requirement for each admitted episode of inpatient rehabilitation since April 2013. In fact, many Level 2b services also now complete the full dataset.

Registration with UKROC requires the submission of:

a) A service profile (updated annually) listing:
   - Capacity (number of beds)
   - Activity levels (occupied bed days per year)
   - Staffing levels (WTE for each discipline and grade)
   - Facilities

b) Parallel ratings of the RCS-E, NPDS and NPTDA for all patients on the unit in cross-sectional tranches at approximately fortnightly intervals (minimum n=100 data pairs).

# Ethics

In 2006, the Ad Hoc Advisory Group on the Operation of NHS Research Ethics Committees prepared the guidelines in the form of *Table 1* below.

The HRA was established in December 2011 to protect and promote the interests of patients and the public in health research, and to streamline the regulation of research. In 2013 it published a document entitled "Defining Research" which took forward the above categories. (http://www.hra.nhs.uk/documents/2013/09/defining-research.pdf)

**Table 1: Research ethics guidelines**

| | Research | Audit | Service evaluation* |
|---|---|---|---|
| 1. | Attempt to derive new knowledge, to generate and test hypotheses. | Designed and conducted to inform delivery of best care. | Designed and conducted to judge current care. |
| 2. | Quantitative research designed to test a hypothesis.<br>Qualitative research to identify/ explore themes following established qualitative methodology. | Designed to answer: "Does this service reach a pre-determined standard?" | Designed to answer: "What standard does this service achieve?" |
| 3. | Address clearly defined research questions aims and objectives. | Measures against a standard. | Measures current service without reference to a standard. |
| 4. | Involves evaluating or comparing interventions especially new ones. Or studying how interventions are experienced. | Involves an intervention in use only. | Involves an intervention in use only. The choice of treatment is that of the clinician and patient according to guidance, professional standards and/or patient preference. |
| 5. | Involves collecting data that are additional to those used for routine care. | Involves analysis of existing /routine data only – but may include administration of a simple interview or questionnaire. | Involves analysis of existing /routine data only – but may include administration of a simple interview or questionnaire. |
| 6. | May involve allocation of patients to intervention groups. Qualitative research uses clearly defined conceptual / theoretical framework. | No allocation to intervention groups- the patient/clinician have chosen the intervention before audit. | No allocation to intervention groups – the patient / clinician have chosen the intervention before evaluation. |
| 7. | May involve randomisation. | No randomisation. | No randomisation. |
| | **ALTHOUGH ANY OF THESE MAY RAISE ETHICAL ISSUES, UNDER CURRENT GUIDANCE:** | | |
| | **REC review required** | **No REC review required** | **No REC review required** |

**Service development and quality improvement may fall into this category.*

At the outset of this programme we sought advice from our local NRES Committee regarding the classification of this programme. It was agreed that it constituted a service evaluation on the basis of the following:

The aims of the programme were to describe and judge current care and, based on these findings, to inform best (including most cost-efficient) care for the future.

- This programme was designated a PbR Improvement Project. The routine collection of the UKROC dataset to record needs inputs and outcomes is a mandated requirement for reimbursement under the Payment by Results.
- All units wishing to qualify for designation as Level 1 or 2 services were required to collect and report the dataset for all case episodes. The data were also used for quality evaluation and benchmarking of services as well as for defining tariff costs.

- Complete collection of data is essential for accurate evaluation – only de-identified data were collated and analysed within the dataset.
- Similarly, the development of protocols for collecting cost data at service or case-episode level are part of required reporting and financial governance within the NHS.

Importantly the programme involved only interventions that were already in use and did not involve randomisation or allocation of patients to treatment groups.

Any descriptive or comparative analysis of these routinely collected activity, costing or outcome data to compare service costs or quality of outcomes do not require REC review or consent from patients or individual staff. Similarly research limited to the involvement of NHS or social care staff recruited as research participants by virtue of their professional role does not require ethics permission.

On the other hand, service evaluations may still be eligible for NHS Support funding under the Comprehensive Local Research Networks (CLRN) programme. Accordingly reporting of patient episodes to the UKROC database was registered with CLRN.

The only aspect of the programme that was identified as requiring ethics permission was the initial testing and psychometric validation of the various tools within the UKROC dataset. Even though these were conducted as secondary analyses of data collected in the course of routine clinical practice, we wished to be able to present them for publication as research articles to enhance the impact of the programme.

The Regional Rehabilitation Unit at Northwick Park Hospital (which houses the UKROC database) has an established track record in tool development and the use of clinical data for research. Since 2002, the unit has held Research Ethics permission from the Harrow Research Ethics Committee (04/Q0405/947) to use, in de-identified form, any of the data collected in the course of routine clinical practice for the purpose of research. The rapid timescale for progression within the earlier stages of this programme did not allow for multiple research ethics permissions to be obtained for each small analysis. Therefore, for many of the tools tested in Chapter 4, the initial evaluation was conducted in this one centre where the requisite ethics permission was already in place. Subsequently the UK Health Research Authority has determined that, *"Research limited to secondary use of information previously collected in the course of normal care (without an intention to use it for research at the time of collection) is generally excluded from REC review, provided that the patients or service users are not identifiable to the research team carrying out the research".*[60]

# Data management.

This is a critical aspect of the programme. Throughout the programme and the research team has been cognisant of the following important issues in relation to data management. These include:

1. Information governance, confidentiality, security and data protection issues for all clinical data, including non-identifiable data.
2. The assembly of a coherent dataset that provides the critical data items to answer the research questions, whilst avoiding undue burden on clinicians.
3. The accuracy and consistency of the data collection – for example ensuring that clinical staff are properly trained in the application of tools and measures, and that these are applied consistently across centres.
4. The integrity of data entry – again ensuring that staff are properly trained and that front-end software supports accurate data entry, for example by denying invalid or inconsistent values.
5. Cleaning and validation of the data prior to analysis.
6. Bringing together the various different data elements for the programme that need to be combined e.g. costing and activity data, with service information and health economics data.

Active involvement of the NHS Information Centre, and subsequently the HSCIC in development of the UKROC dataset ensured that the database met all the national standards for security and protection, and kept up with any changing requirements with respect to information governance.

Collaboration with the Australian Rehabilitation Outcomes Centre was particularly helpful as they had already tackled all of the above areas, from the ethical and strategic issues of data sharing and custodianship, down to on-the-ground data handling and analysis. Involvement of these applicants throughout the project ensured that we continued to build on this experience.

Data management and handling conformed with current national data management guidelines including the Data Protection Act 1998 and NHS Information Governance: Guidance on Legal and Professional Obligations (DH 2007) and NHSE Information Governance Policy 2013.[61]

The UKROC database is held on dedicated NHS network servers at Northwick Park Hospital.

- The Trust has in place full system level security policies to avoid loss of data, covering information and network security.
- The servers are firewall protected and stored in a secure facility with controlled access, on-site security staff and CCTV monitoring.
- Entry to secure areas is restricted to those whose job requires it.
- Data are only transferred by the secure NHS N3 network.

All portable computing devices connected to the Trust network are configured to automatically encrypt any data stored on the hard disk of the computer.

With regard to assuring data quality and consistency, the key requirements were:

- **Training of clinical staff in the various centres to supply data:** This was managed by a series of national training workshops and road shows led by the UKROC team at Northwick Park.
- **Training and supervision/support of research staff to gather and collate data:** All researchers received formal training in Information Governance and kept this knowledge current by regularly updating their training throughout the programme.
- **Careful monitoring to ensure completeness and quality of data collection:** was achieved through routine iterative quality analyses of data to ensure validity, completeness and consistency, both within and between the different data collections.

**A Data Management Group** was established early in the programme to oversee the key aspects of data handling, management and analysis. This was led by the Database Manager (Keith Sephton), supported by Clinical Data analyst (Heather Williams) and programme accountant (Alan Bill). It involved direct input from all researchers as the programme progressed, with input as appropriate from the other key applicants (in particular, RD, RJS and Paul Bassett (Statistician). Meetings reviewed both (a) the research data (parts 1-4) and (b) the case episode data and national database (part 5). The researchers reported on progress with recruitment and data collection, and presented preliminary data summaries, including a running analysis of missing data.

With regard to the national dataset, the aim to achieve a 'workable dataset'; that is reliable with any 'superfluous fields' identified and removed, was considered by the committee throughout the programme, as the database developed. Feedback was obtained throughout the programme through regular contact with the providers during road shows and as they submitted the data.

# Chapter 4: Measuring rehabilitation needs and complexity – tool development

## Objective

To further develop and validate a set of standardised tools to measure individual rehabilitation needs and case-load complexity in neurorehabilitation settings.

**Key deliverables**

1. A set of psychometrically robust tools to measure needs inputs and outcomes that are fit for incorporation into the UKROC database.
2. The tools to be adapted and customised where necessary for the various settings in which they will be used.

## Introduction

As noted in preceding chapters, the UKROC dataset encompasses a set of tools to measure needs, inputs and outcomes for patients with complex rehabilitation needs undergoing specialist in-patient rehabilitation.

This chapter describes the steps that we took to refine and evaluate the set of standardised tools for inclusion within the UKROC dataset.

## Overview of Methods

### *Development and adaptation*

In the course of the programme, we worked with rehabilitation professionals through a series of workshops to identify perceived gaps in the tools and develop items to capture any additional rehabilitation needs/activities that were not already included. This work involved clinicians from a variety of different settings including Level 1 and 2 specialist rehabilitation services, children's services, cognitive behavioural units and specialist nursing homes.

This tool development has been an iterative process that continued throughout the 5-year period, as the instruments were applied, evaluated, and refined to meet the needs of the different rehabilitation services and populations. We acknowledge that, in an ideal world, the tools should have been developed first and applied thereafter. However, the timescale for introduction of PbR did not allow for this orderly progression. For this programme to provide useful data within a 3-4 year time-frame to inform tariffs, it was necessary to progress tool development and data collection in parallel.

## *Settings*

The UKROC database is housed at Northwick Park, which is also the site of the lead clinical service in which many of the tools were first developed. The Regional Rehabilitation Service at Northwick Park includes both inpatient and outreach specialist rehabilitation teams, serving a catchment area that extends across London and the home counties. The breadth of the service and early uptake of the tools into clinical practice, made this an appropriate site for preliminary evaluation as part of the iterative process of development. However, as the tools stabilised and were taken up elsewhere, this provided the opportunity for further multicentre evaluations.

## *Psychometric testing*

### General approach

The data collected through the routine use of these instruments form the basis of a system for describing case-mix, calculating individual patient-level costs and formulating funding models that underpin tariff development for specialist rehabilitation under Payment by Results. Consequently it is imperative that all the measures employed as part of the UKROC database are demonstrated to be psychometrically sound, and that we understand the extent to which they conform to scientific measurement principles as applied to questionnaires and rating scales. Psychometric evaluation formed a substantial underpinning thread to the programme, and it is therefore appropriate to lay out our general approach to this aspect, which was overseen by Professor Richard Siegert.

Outcome tools that are developed within clinical settings typically include a diverse set of items that reflect the multidimensional experience of living with disability. For the most part they generate ordinal data – that is, the numbers assigned to a particular score represent an order, but do not have true numerical value. Although the scores are often added together to a single overall value, the question of validity arises when summating ordinal, as opposed to interval, data.

Psychometric evaluation provides an assessment of the extent to which a tool provides a valid and reliable measure of the trait that it purports to capture. Tools designed to capture a broad concept such as 'disability' will often require reduction into several subscales or 'constructs' that form valid and reliable sub-scales to measure different elements of progress that an individual or group of patients may make during rehabilitation. The creation of a scale with sound measurement properties may involve discarding items that do not fit the mathematical scaling model, but sometimes it is these mis-fitting items that provide the most critical information for clinical decision-making. Therefore, there is sometimes a tension between the requirement for optimum measurement properties within a scale and its ability to capture all the clinical information that is most relevant to patients, families and their treating teams.

The term 'clinimetrics' was popularised in the 1980s by Alvan Feinstein, an American physician who argued that, first and foremost, clinical measures must provide the information that is important to clinicians caring for patients.[62] This is not to suggest that their statistical properties are unimportant, but rather that measures must start by providing answers to important clinical questions while being brief and simple to administer. In the language of psychometrics this means the instrument must have '*face validity*' and '*utility*'.

Even within the field of psychometrics there are a number of different schools of thought, often generating quite polarised opinion, about the best approaches to evaluation of measurement properties. For example:

- *classical test theory* involves the identification of individual constructs within a scale that allow grouping of items into factors that form a valid and reliable ordinal scale for measurement within the tool. Commonly used approaches include exploratory factor analysis (e.g. using principal component analysis (PCA) followed by confirmatory factor analysis to confirm the 'goodness of fit' to the model.

- *item response theory* takes account of both the ability of the individual and the difficulty of the item to develop a mathematical model that has scaling properties that approximate to interval quality data. Commonly used approaches include Mokken scaling (non-parametric) and Rasch Analysis (parametric).

Patients with complex neurological disability typically present with a wide range of impairments including physical, cognitive, emotional, behavioural and psychological disabilities. Tools that capture this range of experience will typically be multi-dimensional and frequently non-linear in nature. In addition, the measures developed for the UKROC database had to be fit for use in routine clinical practice within the modern NHS. In other words, notwithstanding their psychometric merits, they need to be measures that clinicians will embrace in their day-to-day work with patients. They must be timely to administer and provide information that has practical application to inform clinical decision-making. At the same time the UKROC measures form part of a national dataset and will certainly be used for research and audit purposes. Hence it is essential that while meeting the clinimetric criteria for a good measure they also have robust enough psychometric properties for inclusion in statistical analyses.

Consequently we adopted a conservative but rigorous strategy to guide the development of measures in the present project. Our approach to developing and evaluating the UKROC measures combined traditional test theory statistical methods with an emphasis on clinimetrics. We used the broad range of statistical methods subsumed under the rubric of 'classical test theory'.

Nevertheless, once the clinimetric and psychometric properties of a clinical measure are known, it may become relevant to explore the extent to which our more commonly used scales may provide interval-quality data. Some of our instruments provide this through intuitively-developed computerised algorithms – for example the Northwick Park nursing and therapy Dependency Scales[49, 53] translate into an estimation of care/staffing hours and the cost of providing the appropriate care package in the community to fit the timetable of care needs.[51] In this case, conversion to interval quality data through theoretical mathematical modelling becomes superfluous. In other instances it is pertinent to explore the extent to which a scale fits an interval level scale using a Rasch modelling approach or Mokken Analysis, which we have also undertaken in some instances.

## Psychometric methods

The general criteria for a sound psychometric instrument are well known. We used the framework developed by the Scientific Advisory Committee of the Medical Outcomes Trust (Terwee et al. 2007)[63] in *Table 2* below as the basis for reporting the psychometric properties of the tools. In addition to the Terwee criteria, there is an important clinimetric property namely that clinicians must *want* to use the tool. We have therefore added a new criterion which we have called 'Engagement'.

**Table 2: Medical Outcomes Trust framework for psychometric evaluation**

| Attribute | Criteria | Evaluation |
|---|---|---|
| **Conceptual and measurement model** | ***The rationale for and description of the concept and the populations that the measure is intended to assess*** | |
| | **Clinical content and design** | The concept to be measured, and the intended target population (who should be involved in scale development)<br>The structure of the scale and intended level of measurement are clearly described and there is an empirical basis for selection of the clinical content and item combinations. |
| | **Dimensionality** | Information on dimensionality and rationale for deriving scale scores<br>Any intended subscales are built into the conceptual design<br>Exploratory and Confirmatory Factor Analyses within target population to confirm the factor structure |
| **Reliability** | ***The degree to which the instrument is free from random error*** | |
| | **Internal consistency** | The precision of the scale based on homogeneity of the scales items at one point in time<br>Item-total correlations (correlation of each item with the total score (excluding that item)<br>Cronbach's alpha calculated for each subscale – ideally between 0.7-0.95 |
| | **Reproducibility** | Stability of the test over time (repeatability, intra-rater reliability) and between different observers (inter-rater reliability)<br>(Time period should be long enough to prevent recall, but short enough that no clinical change has occurred – describe and justify)<br>Total scores – intra-class correlations (ICC) (≥0.7) or Bland and Altman Limits of agreement for continuous (interval) data<br>Item-by item analysis of agreement using weighted kappa statistics |
| **Validity** | ***The degree to which the instrument measures what it purports to measure*** | |
| | **Content** | Confirmation that the domain (s) of the instrument are appropriate to its intended use<br>E.g. Feedback from clinicians and target population that items of interest are comprehensively represented |
| | **Criterion-related** | Evidence that the scores of the instrument are related to an accepted gold standard<br>Where a gold standard is available |
| | **Construct-related** | Evidence that the scores relate to other instruments consistent with theoretically-driven hypotheses<br>The extent to which the scores correlate with other measures of similar concepts (concurrent validity) and do not correlate with unrelated measures (discriminant validity) |

| Attribute | Criteria | Evaluation |
|---|---|---|
| **Responsiveness** | ***Ability to detect change over time where real changes occur*** | |
| | **Change: admission to discharge** | Evidence of change in longitudinal analysis – significant differences (Wilcoxon z or paired T tests), Effect size estimation. |
| | | Ability to distinguish clinically important change from measurement error<br>(e.g. Limits of Agreement should be smaller than Minimal Important Change (MIC)* ; or Gyatt's Responsiveness Ratio: MIC/SDC = at least 1.96 for continuous normally-distributed data). |
| | **Floor/ceiling effects** | No significant floor or ceiling effects are present<br>(<15% of subjects achieve the highest or lowest possible score in a sample size of at least 50 subjects in the target population) |
| **Interpretability** | ***The degree to which easily understood meaning can be assigned to the quantitative scores*** | |
| | **Clinical meaning** | Describe how the tool should be reported and interpreted e.g. sub-scores, total scores etc<br>Minimal Important Change (MIC) is defined for the target population<br>Comparison of change in groups expected to change with those not (e.g. responders versus non-responders, active treatment versus placebo) |
| **Burden** | ***The time, effort or other demands of administering the instrument*** | |
| | **Time to administer** | Information on average and range of time taken to complete the instrument.<br>Any special requirements or restrictions – e.g. training, level of professional expertise to apply it<br>Under what circumstances is it not suitable? |
| **Engagement** | ***Clinicians want to use the scale as they value the information it provides*** | |
| | **Clinician engagement** | Training tools / accreditation system provided to ensure accuracy of data collection<br>User-friendly software is available to embed the use of the tools into the everyday clinical decision-making.<br>There is backing from senior management to support use of the tool in routine clinical practice |
| **Alternative modes of administration** | | E.g. clinician administered / self-complete etc. Describe if these are available<br>Evidence of reliability, validity, responsiveness, interpretability and burden for each administration<br>Information on comparability of the alternative modes |
| **Cultural and language adaptations** | | Describe if these are available<br>Methods to achieve conceptual and linguistic equivalence.<br>Any significant differences between the original and translated versions – how any differences were reconciled |

**Smallest Detectable Change SDC=1.96 x √2 x Standard Error of Measurement (SEM); SEM = √error variance of an ANOVA analysis*
*Minimal Important Change (MIC) = "the smallest difference which patients would perceive as beneficial and would mandate an change in the patients management (in the absence of troublesome side effects and excessive cost")*

## Explanatory note for figures and graphs

### *Boxplots*

When boxplots are presented, 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.

### *Path diagram for confirmatory factor analyses*

Circles (or ellipses) represent unobserved latent variables, squares (or rectangles) represent observed variables and single arrows represent the impact of one variable on another. A double arrow represents the correlation coefficient between two variables. "e" enclosed in a circle indicates error – representing measurement error. The single-headed arrows represent linear dependencies. The numeric values located with single-headed arrows give an estimate of standardised regression weight (standardised maximum likelihood parameter).

## Dimensionality

Clinical tools used in the context of complex disability are rarely uni-dimensional. A key task for psychometric analysis is to determine whether it is still reasonable to sum the items to a single total score, or whether the instrument should be divided into two or more subscales for the purpose of analysis.

'Dimensionality' refers to the number of distinct constructs that can be derived from a single questionnaire or rating scale. We used exploratory factor analysis (EFA) for determining the dimensionality of all newly developed scales and confirmatory factor analysis (CFA) to confirm the structure identified.

- For EFA we aimed to use at least ten participants per item as the minimum sample size, and we applied principal component analysis (PCA) with Varimax rotation as these typically provide clear, interpretable solutions. In all cases, we used the Keyser Myer Olkin test and Bartlett's Test of Sphericity to ensure that the correlation matrix was suitable for factor analysis. The decision as to the number of factors to rotate was based upon consideration of (a) the number of factors with eigenvalues >1, (b) Visual inspection of the scree plot, and (c) parallel analysis.
- CFA was completed on a separate sample using the IBM-SPSS AMOS software. The fit statistics we used included: Chi-square/df, Comparative Fit Index, Goodness of Fit Index, Root Mean Square Error of Approximation (RMSEA). Wherever possible, we used 200+ participants as the minimum sample size for CFA.
- In tools where the data were highly skewed, we also used Mokken scaling, a non-parametric item response model to confirm the scale was a reliable, unidimensional ordinal scale. Mokken scaling analysis of polytomous items was undertaken using MSPWIN 5.0 software.[64]

Unlike parametric methods, such as confirmatory factor analysis or Rasch,[65-67] Mokken analysis makes no assumptions concerning the distribution underpinning the data. It calculates Loevinger's H Coefficient for a scale, and each of its individual items, to determine if they satisfy the requirements for a stochastic or probabilistic Guttman scale.[68] H values < 0.30 are considered to reflect a poor item and values for the scale as a whole are interpreted as follows: H values in the range 0.30 - 0.40 reflect a *weak* scale, 0.40 – 0.50 a *medium* scale and H>0.50 reflects a *strong* scale. Hence any scale that fulfils the criteria for a robust scale in Mokken terms can be considered a reliable, uni-dimensional, ordinal scale that is suitable for rank-ordering persons.

## Reliability

Scales should be internally consistent and give reproducible results.

The most widely used index of **internal consistency** is Cronbach's alpha, which should be between 0.70 – 0.95 for a scale to have good internal consistency. Cronbach's alpha was calculated for all subscale and total scales.

**Reproducibility refers to the stability of test scores over a fairly short interval of time during which the** patient is unlikely to have changed markedly, or between observers. Broadly speaking there are three types of reproducibility:

Test-retest reliability is used for patient self-report questionnaires. Typically this involves a sample of patients completing the same questionnaire (e.g. a depression symptom inventory) several days apart and calculating the correlation between the total scores on the two occasions;

Intra-rater reliability is tested when a clinician rates the same patients on two separate occasions. The interval between ratings should be long enough to prevent the clinician simply remembering their previous scores, but brief enough so that any patient's condition is unlikely to have changed significantly;
Inter-rater reliability is tested when different clinicians or clinical teams rate the same patient independently of each other.

We examined the extent of agreement between item scores both between and within raters. Intraclass correlation coefficients (ICCs) were calculated for subscale and total scores using SPSS (IBM). For individual item scores, we calculated absolute agreement and weighted kappa coefficients using STATA (Statcorp). These were interpreted according to Landis and Koch[69] (kw 0.01-0.2 = Slight agreement; kw 0.21-0.4 = Fair agreement; kw 0.41-0.60 = Moderate agreement; kw 0.61-0.80 = Substantial agreement; kw 0.81-1.0 = Almost perfect agreement.

## Validity

Validity refers to the extent to which a test can be shown to measure the construct that it purports to measure. It is often considered in terms of Content, Criterion and Construct validity.

- **Content validity** means that a measure assesses all the important elements that constitute the domain of interest and no irrelevant items. In this programme, content validity was typically established through an interactive process of pilot application in clinical practice and feedback from workshops and other interaction with clinicians and/or patients, to obtain their views on the relevance and applicability of items to their particular condition/setting.
- **Criterion validity** involves comparing scores on a measure with an existing 'gold standard' where one exists. A problem for most of the clinical measures in this programme was the lack of a gold standard. In practice this meant that criterion validity was most often demonstrated through correlations with other scales that measure related constructs.
- **Construct validity** The extent to which the scores correlate with other measures of similar concepts (concurrent validity) and do not correlate with unrelated measures (discriminant validity). These were tested in a hypothesis-driven manner through examination of correlations with other instruments rated in parallel. Spearman rank correlations were used for short-ordinal data (i.e. ordinal scores with few scoring levels) – or where data were highly skewed. Pearson correlations for interval quality data (or long-ordinal data that were distributed within acceptable limits of normality.

## Responsiveness

Responsiveness concerns the ability of a measure to detect real change in the status of a patient on the phenomenon it measures. For an outcome measure it is typically tested by examining change between the start and end of the programme. We demonstrated the responsiveness of our measures by comparing the mean or median scores for a large sample of patients at admission to neurorehabilitation and at discharge. Where the sample size was large and approximately normally distributed we used parametric statistics (e.g. paired t tests) wherever possible. Responsiveness was calculated using Kazis' Effect Size (mean change from baseline/SD baseline)[70] and interpreted according to Cohen[71] (i.e. 0.2 = small, 0.5 = moderate, 0.8 = large). Where smaller numbers were involved, or the distribution was substantially skewed, we used non-parametric tests (e.g. Wilcoxon paired rank tests) or parametric tests with boot strapping to minimise bias.

## Interpretability

Where appropriate, the minimal important change (MIC) was examined. However, for multidimensional scales, different items will impact more or less on the outcomes of interest – for example within the FIM score, and changes in continence, or items that require intervention many times a day will have a greater impact on overall care needs than others. We therefore explored other ways to ensure interpretability, such as the use of 'FAM-splats' to provide clinicians with a 'view at a glance' impression of the changes in item scores within the scale.

## Utility

Utility essentially reflects the balance between the benefits that derive from using an instrument and the burden of applying the instrument. For instruments to be usable in the course of routine practice, **rater burden** is a key consideration – for example it is important to know how long it takes to administer. On the other hand, clinicians may be prepared to put in extra time and effort if they find the end result worthwhile – for example, the 30-item UK FIM+FAM clearly takes longer to rate than the 10 item Barthel Index, but in the context of complex disability (where cognitive and psychosocial problems are often the major limiting factors), clinicians are often prepared to spend the extra time to rate a measure that provides a more holistic evaluation of the key issues. In this programme, utility was assessed through feedback from clinicians, but also by the completeness of data.

## Engagement

Principal requirements for the successful integration of data gathering are a) staff training and b) on-going commitment from the senior staff and managers. These are both essential to the maintenance of effort required to provide assurance of data quality in the longer term. But more than that, if a tool is to be recorded consistently in routine practice, it is critical that clinicians feel clinically engaged in the process and actually <u>want</u> the information. In the course of this programme we have found that key steps to clinician engagement have been the development of user-friendly software and embedding the use of the tools into the everyday clinical decision-making.[72] As the UKROC software was developed, we took care to include automated features, such as production of the 'FAM-Splat' and the care needs time table within the Northwick Park Dependency Scale, that support decision-making regarding the planning and provision of care at the point of transfer back into the community.

## Alternative modes of administration

While the majority of instruments used in a clinical setting are traditionally administered by the treating team, patient and family perspectives are increasingly receiving greater emphasis. Moreover, follow-up after the patient leaves in-patient rehabilitation often relies on self-report methods such as postal/online questionnaires and or telephone follow-up. An important aspect of this programme was to be able to evaluate patients in a variety of different settings, so self-report versions have been generated for some of the tools. In this case it is important to test how far the self-reported and clinician-reported versions produce the same information.  A perfect match is not always expected however as patients and clinicians are likely to have different perspectives, and in some instances this difference itself may be the target of enquiry.

**Cultural and language adaptations**

Many tools that evaluate activities and participation are sensitive to different cultures and, in mixed communities, it is sometimes necessary to deliver questionnaires in different languages.  As this programme was administered in the UK, and mainly in the context of in-patient or residential rehabilitation, the production of tools in different languages has not been a key focus for this research. Nevertheless, some of the UKROC tools have already been taken up internationally and translated into other languages, so the presence of these translations was noted where relevant.

# Measuring the complexity of rehabilitation needs – the Rehabilitation Complexity Scale (RCS)

## *Background*

The Rehabilitation Complexity Scale (RCS) was designed to provide a simple measure of the complexity of rehabilitation needs and/or interventions (i.e. resource use), which would ultimately form the  measure for specialist rehabilitation services.

Apart from length of stay, the key factors that determine the costs of rehabilitation for a given individual are their needs for care, nursing, therapy and medical intervention.  The RCS was developed as a simple measure of these parameters, which is timely to apply in a clinical setting. The item levels are broadly ordinal, but were designed to reflect clinically important features that may influence care planning, rather than being chosen for their interval qualities.  Item scores are expected to be in some way cumulative, but it is pertinent to know whether they may reasonably be 'summed' to give a total score as an overall indicator of caseload complexity; and if not, whether they can be grouped, or should be reported individually.

Prior to the start of the programme, the basic RCS tool (RCS-version I) had been published, together with a preliminary exploration of  its validity in distinguishing differences in caseload complexity across a range of specialist rehabilitation services, and of its practicality for routine use for this purpose in routine clinical practice.[16]   In a multi-centre cross-sectional analysis of 677 ratings from 49 services it showed clear differences between tertiary ('complex specialised') and secondary ('district specialist') rehabilitation services, on the basis of their relative proportions of complex cases.  However, the therapy subscale, which recorded total hours of therapy intervention, was found to be difficult to rate prospectively.

The instrument was subsequently revised through several iterations to form the RCS version 8 (also known as the RCS II), in which the care, nursing and medical scales remained the same, but the therapy scale was divided into two subscales reflecting (a) the number of therapy disciplines and (b) the overall intensity of treatment.  Throughout the course of this programme it has continued to evolve through an iterative process of development and evaluation.  The tool has proved to be a robust and practical measure of caseload complexity, but the evaluations have led to its refinement and adjustment for customised use in a number of different settings, as described below.  *Table 3* summarises the item content and score ranges of the various versions.

**Table 3: Comparative content and score ranges of the RCS-v1, RCS-v8 and RCS-E versions 12 and 13**

| RCS original (v1) | | | Subsequent versions | | RCS v8 | RCS-E v12 | RCS-E v13 | RCS-SNH |
|---|---|---|---|---|---|---|---|---|
| Item | | Range | Item | | Range | Range | Range | Range |
| C | Basic care | 0-3 | C/R | Basic care/or Risk | 0-3 | 0-4 | 0-4 | 0-3 |
| N | Skilled nursing needs | 0-3 | N | Skilled nursing needs | 0-3 | 0-3 | 0-4 | 0-3 |
| T | Therapy | 0-6 | TD | Therapy disciplines | 0-3 | 0-4 | 0-4 | 0-3 |
| | | | TI | Therapy intensity | 0-3 | 0-4 | 0-4 | 0-3 |
| M | Medical needs | 0-3 | M | Medical needs | 0-3 | 0-3 | 0-4 | 0-3 |
| | | | E | Equipment/facilities | - | 0-2 | 0-2 | 0-3 |
| **Total Range** | | **0-15** | | | **0-15** | **0-20** | **0-22** | **0-18** |

## *RCS II (version 8) - Initial clinimetric evaluation*

The first formal evaluation of the RCS within this programme was a clinimetric evaluation of the RCS II (known as the RCS-v8).[47] We used the Medical Outcomes Trust framework described by Terwee et al 2007[63] to summarise the evidence for its clinimetric properties.

The evaluation was based on an observational cohort analysis of 179 consecutive patients (mean age 44.5 (sd 15) years, Males:Females 110:69) with complex neurological disabilities, mainly following acquired brain injury, undergoing in-patient rehabilitation in a single tertiary specialist setting. All data were collected in the course of routine clinical practice. Data were entered at the time of collection into the dedicated UKROC software. The methods of analysis followed the principles laid out earlier in this chapter. The paper giving the full methodology, is published in open access in the Journal of Neurology, Neurosurgery and Psychiatry.[47] As this is available in the public domain, only the key findings will be summarised here.

### Dimensionality

An exploratory factor analysis demonstrated that all five items of the RCS v2 loaded 'moderate' to 'high' on the first un-rotated principal component with loadings ranging from 0.52-0.79 and a Cronbach's $\alpha$ co-efficient for the total scale of 0.76. Principal components analysis (PCA) with varimax rotation suggested a two-factor solution accounting for 68% of the total variance in scores:

- The first factor 'Nursing/medical' care accounted for 44% of the variance. The C, N and M items all loaded high (0.65-0.88) on this factor and low (<0.15) on factor 2.
- The second factor 'Therapy', accounted for 24% of the variance. The two therapy items (TD and TI) both loaded above 0.80 on this factor and low on factor 1.

Confirmatory factor analysis, conducted on RCS scores at discharge (n=173), confirmed that this two-factor model had an excellent fit with a Goodness of fit index (GFI) = 1.0, Comparative Fit Index (CFI) = 1.0, an Adjusted Goodness of Fit Index (AGFI) = .98, and Chi-square = 1.92, (df=4, p=0.750).

### Reproducibility (test-re-test repeatability)

Quadratic weighted kappa coefficients for agreement between repeated RCS ratings for the various items 0.69-0.96,which constitutes 'substantial to 'almost perfect' agreement according to the interpretation of Landis and Koch [69]

### Validity

In the absence of a gold standard, convergent and discriminant validity were examined with reference to parallel application of:

- Northwick Park nursing and therapy dependency scores (measures of rehabilitation inputs), and
- Barthel Index (BI) and Functional Independence Measure (FIM) scores, as global measures of disability and independence for basic activities of daily living.

All of the measures were applied retrospectively in respect of the average levels of intervention provided over the preceding fortnight.

We expected to find a close correlation between the RCS care and nursing (C and N) items and the Nursing Dependency (NPDS); and between the two therapy measures (RCS T score and NPTDA), but weaker correlations across the therapy/nursing divide.

These relationships are summarised in *Table 4*. The RCS total score demonstrated moderately strong correlations with the NPDS and NPTDA (rho 0.49–0.79, p<0.001). However, within the subscales there were differential correlations. Only the RCS-T scale correlated strongly with the NPTDA (rho 0.72), whilst the NPDS correlated strongly with the C, and N items (0.70-0.80), and to a lesser extent with the M item (0.38) but only weakly with the T scale (0.26). These relationships suggest that the RCS-T score reflects the needs for therapy intervention and the RCS-C and -N scores reflect care and nursing needs - but that, as expected, these are relatively independent of each other.

We also expected to find a broad relationship between complexity and global measures of disability – the more disabled patients being likely to have greater needs for care and nursing, and possibly also for therapy. Because BI and FIM are measures of independence (as opposed to dependency), negative correlations with the RCS, NPDS and NPTDA were expected, and indeed found. FIM motor and BI scores correlated strongly with basic care and nursing scores (Spearman rho -0.65 to -0.79), but only weakly with therapy (rho -0.26), except for the FIM Cognitive subscale (rho= -0.44 (p<0.001).

As expected, there was very little relationship between the FIM or Barthel and RCS-Medical scores (rho -0.28 to -0.33) as these global disability measures do not contain any information regarding the patient's medical condition scores. Similarly, as sicker patients are likely to require both medical and nursing care we expected and found a closer relationship between the RCS-M score and nursing needs (rho 0.38), than with therapy needs (rho 0.19).

**Table 4: Spearman correlations between Rehabilitation Complexity Scale items (n=179) and Northwick Park Dependency Scales, Barthel Index and Functional Independence Measure at the start of treatment**

| | Rehabilitation Complexity | | | | | Dependency | |
|---|---|---|---|---|---|---|---|
| | **Care** | **Nursing** | **Therapy TD+TI** | **Medical** | **Total** | **NPDS** | **NPTDA** |
| **RCS[a] Scores** | | | | | | | |
| **Nursing** | **0.64**** | | | | | | |
| **Therapy (TD+TI)** | 0.18 | 0.23 | | | | | |
| **Medical** | 0.32* | **0.44**** | 0.20 | | | | |
| **Total RCS** | **0.72**** | **0.79**** | **0.64**** | **0.64**** | | | |
| **Dependency scores** | | | | | | | |
| **Total NPDS** | **0.80**** | **0.70**** | 0.26 | **0.38**** | **0.73**** | | |
| **Total NPTDA[c]** | 0.16 | 0.18 | **0.72**** | 0.19 | **0.49**** | 0.22 | |
| **Barthel and FIM[d] scores** | | | | | | | |
| **Barthel Index** | **-0.76**** | **-0.65**** | -0.26 | -0.28* | **-0.67**** | **-0.85**** | -0.26 |
| **FIM Motor** | **-0.79**** | **-0.69**** | -0.26 | -0.33* | **-0.72**** | **-0.88**** | -0.26 |
| **FIM Cognitive** | -0.33* | -0.33* | **-0.44**** | -0.17 | **-0.47**** | **-0.47**** | **-0.52**** |

*Spearman Rank Correlation tests: **p<0.001, *p<0.01, (to account for multiple tests the threshold for significance was taken as p<0.01 – values above this are considered non-significant).*
*[a] Rehabilitation Complexity Scale: Subscales: C = Care, N=Nursing, T=Therapy, M=Medical*
*[b] NPDS: Northwick Park nursing Dependency Scale, [c] NPTDA: Northwick Park Therapy Dependency Assessment*
*[d] FIM: Functional Independence Measure (Motor and Cognitive subscales)*

### Responsiveness

We did not expect the RCS overall to change markedly during the programme, even though we anticipated that the relative components of care/nursing and therapy might change. Overall there was a small, but significant, reduction in RCS - principally reflecting the reduction in care, nursing and medical (C+N+M) needs (Wilcoxon z -9.0 p<0.001) - whilst the therapy component (TD+TI) increased overall (Wilcoxon z -4.6 p<0.001).

*Figure 6* shows an example single case analysis of serial RCS measurements at fortnightly intervals, during a five-month rehabilitation programme. Care and nursing needs gradually fell during the stay, but therapy needs followed a variable course as the interventions changed at different stages of the programme. Similar patterns were reflected also in the NPTDA/NPDS scores and hours of intervention.

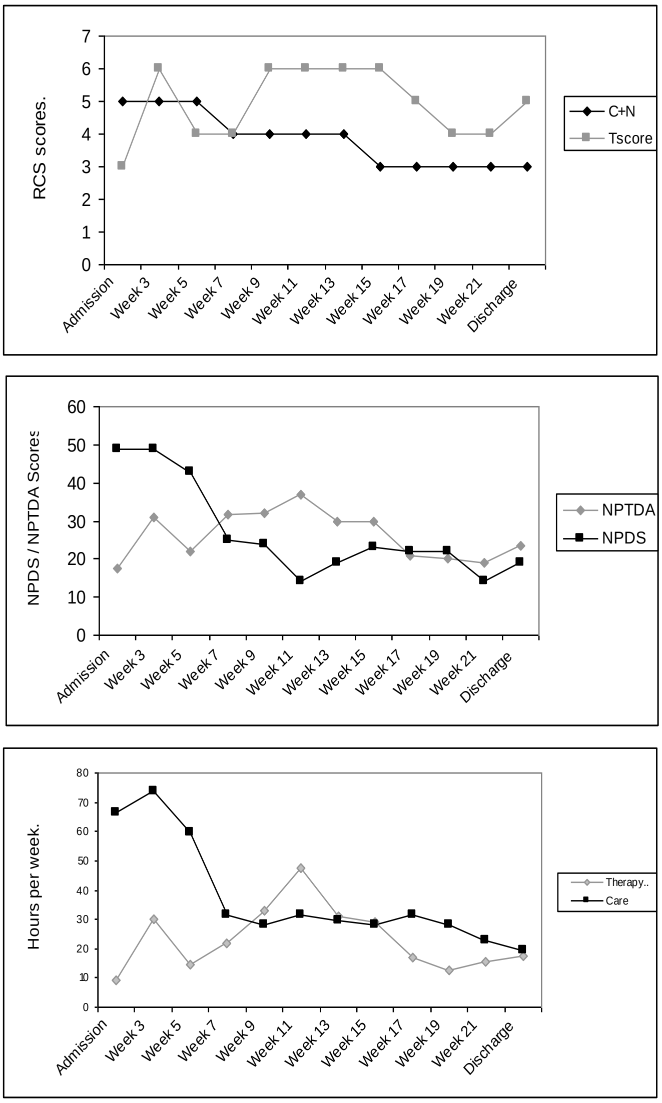



**Figure 6: Serial RCS scores in a single case episode of approximately 5 months stay compared with NPDS/ NPTDA scores, and care and therapy hours per week for the equivalent time points**

### Feasibility and clinical feedback

In this study, the RCS took less than 1-2 minutes to administer by a team who was familiar with the scoring manual and used the score regularly in routine practice.

Clinicians working in tertiary specialist neuro-rehabilitation settings reported favourably on utility, content and face value, but recorded some further problems.

a) Ceiling effects were noted for patients with very complex needs – particularly within the therapy subscales.

b) Needs for special equipment / facilities were not identified.

c) The Care section did not capture the 'risk' or needs for supervision of patients who were ambulant but confused, for example in cognitive behavioural rehabilitation settings.

### Limitations of this study were:

- It was conducted on a cohort from a single centre with a particularly complex group of patients undergoing neurological rehabilitation.
- Our confirmatory factor analysis was undertaken on discharge scores from the same group of patients as the exploratory factor analysis.

### Overall summary of clinimetric properties

An overall summary of the clinimetric properties of the RCS v 8 according to the Medical Outcomes Trust framework is shown in *Table 5*.

**Table 5: Summary of clinimetric properties of the Rehabilitation Complexity Scale v8**

| Attribute and criteria | Evaluation |
|---|---|
| **Conceptual and measurement model:** ***The rationale for and description of the concept and the populations that the measure is intended to assess*** | |
| **Clinical content and design** | The Rehabilitation Complexity Scale (RCS) is a five-item ordinal scale, scored on a range of 0-15.<br>Designed to provide a simple measure of the complexity of rehabilitation needs and/or interventions, which is timely to apply and takes account of basic care, specialist nursing, therapy and medical interventions.<br>In this evaluation, it was tested in the context of 'interventions provided'. |
| **Dimensionality** | Exploratory and confirmatory factor analyses showed strong evidence that the RCS has two distinct dimensions ('Nursing/medical care' and 'Therapy'). |
| **Reliability:** ***The degree to which the instrument is free from random error*** | |
| **Internal consistency** | Cronbach's alpha 0.76 and item-total correlations (0.51-0.78) showed moderate internal consistency. |
| **Reproducibility** | Test–retest repeatability after two hours using the ward-round as a distracter task:<br>Quadratic-weighted Kappa values were 0.93, 0.96 and 0.94 for the care, nursing and medical items respectively - constituting excellent agreement. Repeatability for Therapy (TD and TI) items was not tested in this evaluation. |
| **Validity:** ***The degree to which the instrument measures what it purports to measure*** | |
| **Content** | The RCS care (C), nursing (N), therapy (T) and medical (M) items are the principal 'causes' of case complexity, which (together with length of stay) ultimately determine the cost of a rehabilitation episode. |
| **Criterion-related** | Not testable - no accepted gold stand currently exists. |
| **Concurrent** | Convergent and discriminant validity tested against Northwick Park Nursing and Therapy Dependency Scores:<br>The RCS Care/Nursing scores correlated strongly with other measures of nursing dependency (rho 0.70-0.80; p<0.001) whilst the RCS-T scores* correlated weakly (0.26, p non-significant). Conversely, the RCS-T scores correlated strongly with the therapy dependency (0.72, p<0.001). |

| | |
|---|---|
| **Responsiveness:** ***Ability to detect change over time where real changes occur*** | |
| **Change: admission to discharge** | RCS scores changed significantly over the course of a 3-4 month stay but the items changed in different directions:<br>Care, nursing and medical interventions (C+N+M) reduced (Wilcoxon z -9.0 p<0.001); whilst therapy interventions(TD+TI) increased (Wilcoxon z -4.6 p<0.001). |
| **Interpretability:** ***The degree to which easily understood meaning can be assigned to quantitative scores*** | |
| **Clinical meaning** | The RCS is recommended to be reported by item: e.g. RCS 8 = C2 N1 T4 M1, as the level descriptors provide a clinical description of needs/interventions that is useful for treatment planning.<br>This evaluation also supports summation into two subscales: Nursing medical care (C+N+M) and Therapy (TD+TI). |
| **Burden:** ***The time, effort or other demands of administering the instrument*** | |
| **Time to administer** | The RCS is designed to be intuitive and requires minimal training.<br>In this study, it took less than one to two minutes to administer by a team who was familiar with the scoring manual and used the score regularly in routine practice. |
| **Alternative modes of administration** | None currently available. |
| **Cultural and language adaptations** | An Italian version is now available. |

** RCS T score = Therapy Disciplines (TD) +Therapy Intensity (TI)*

### Conclusion from this study:

In this cohort, the RCS provided a reliable, valid and moderately responsive profile of rehabilitation interventions, separating into two main subscales. It usefully identified medical and therapy inputs not captured by the FIM and Barthel Index, which are commonly used (e.g. in the US and Australia) to define case complexity and in rehabilitation. The change in pattern over time mirrored changes in therapy and nursing inputs as measured by the NPDS and NPTDA, suggesting that serial rating of the RCS score may provide a reasonably reflection of variation in the use of resources (staff time) and therefore of changes in the cost of rehabilitation as the patient moves through the programme. However, in view of the identified ceiling effects, we concluded that before the RCS could be used as a case-mix measure for specialist rehabilitation, further refinement was required to ensure that it provided an accurate reflection of resource use, providing this could be achieved without adding significantly to the burden of rating.

## *Further development - The Rehabilitation Complexity Scale extended (RCS-E)*

To address the acknowledged weaknesses, an extended version of the Rehabilitation Complexity Scale (the RCS-E) was subsequently developed and subjected to a comparative evaluation with the RCS-v8. Once again the RCS-E went through an iterative process of development before it stabilised at RCS-E-version 12.

Evaluation of the RCS-E-v12 was conducted on fortnightly ratings recorded over a 40-month period between 6$^{th}$ June 2006 and 7$^{th}$ September 2010 in one tertiary specialist rehabilitation service. The RCS-v8 and the RCS-E were recorded in parallel by trained members of the multi-disciplinary team during routine weekly team meetings. Serial RCS ratings of the level of care, nursing, therapy and medical interventions were examined for dimensionality, repeatability, consistency and responsiveness; and compared with the Northwick Park Dependency Scales (NPDS and NPTDA), and the FIM and Barthel Index, recorded at the start and end of treatment.

All data were collected in the course of routine clinical practice. The methods of analysis followed the principles laid out earlier in this chapter. The paper giving the full methodology, is published in open access in Disability and Rehabilitation.[48] Once again, as this is available in the public domain, only the key findings will be summarised here.

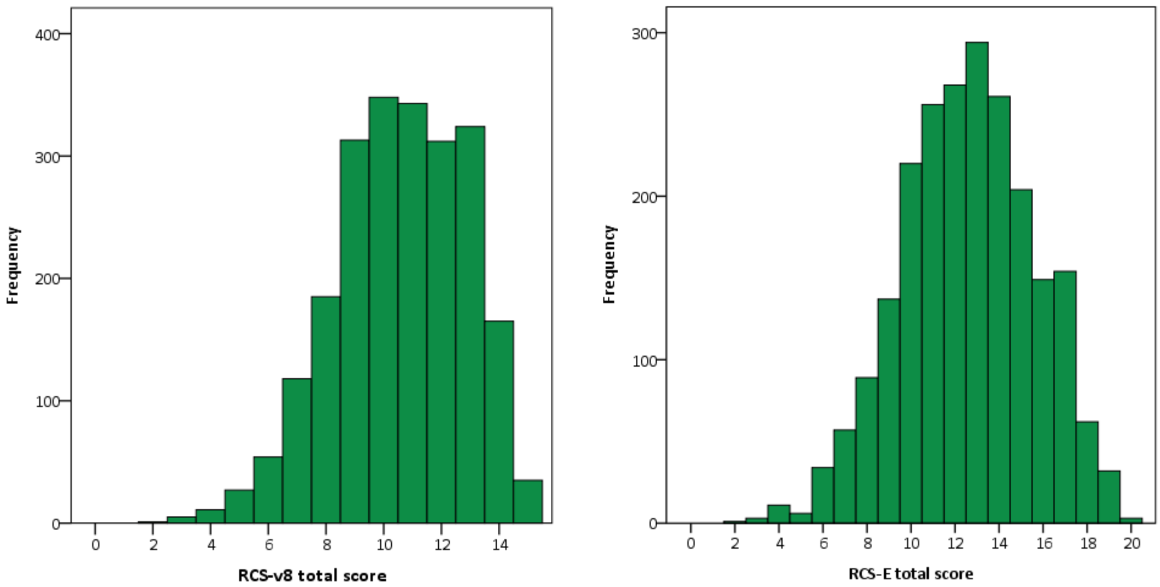
A total of 2241 valid serial ratings were obtained on admission from 331 patients, mean age 44 (SD 15) years, range 15-80, male: female ratio 60:40%. A total of 269 (81%) had acquired brain injury (of which the aetiologies were 61% stroke, 22% traumatic, 6% hypoxic and 11% other (e.g. inflammatory); 20 (9%) had spinal cord injury and 28 (9%) had peripheral neurological conditions e.g. Guillain Barre Syndrome or critical illness neuropathy).

Descriptive statistics for complexity and dependency scores are shown in Table 6. The data represented nearly the full range of both RCS-v8 and RCS-E scores, as shown in *Figures 7a and 7b.* The RCS-E showed greater separation of complexity levels in the upper part of the scale, compared with the RCS-v8, resulting in a more normal pattern of distribution.

**Table 6: Descriptive statistics for Complexity and Dependency ratings across all time points of the sample (n=2241)**

| Rehabilitation Complexity Scores | Mean | (SD) | Median | (IQR) | Range |
|---|---|---|---|---|---|
| **RCS-v8** | | | | | |
| **Care (0-3)** | 1.4 | (0.8) | 1 | (1-2) | 0-3 |
| **Nursing (0-3)** | 2.2 | (0.8) | 2 | (2-3) | 0-3 |
| **Therapy Disciplines (0-3)** | 2.8 | (0.4) | 3 | (3-3) | 1-3 |
| **Therapy Intensity (0-3)** | 2.4 | (0.6) | 2 | (2-3) | 0-3 |
| **Medical (0-3)** | 1.8 | (0.7) | 2 | (1-2) | 0-3 |
| **Total (0-15)** | 10.6 | (2.3) | 11 | (9-12) | 2-15 |
| **RCS-E-v12** | | | | | |
| **Care (0-4)** | 1.4 | (0.9) | 1 | (1-2) | 0-4 |
| **Nursing (0-3)** | 2.2 | (0.8) | 2 | (2-3) | 0-3 |
| **Therapy Disciplines (0-4)** | 3.0 | (0.7) | 3 | (3-4) | 1-4 |
| **Therapy Intensity (0-4)** | 2.6 | (0.8) | 2 | (2-3) | 0-4 |
| **Medical (0-3)** | 1.8 | (0.7) | 2 | (1-2) | 0-3 |
| **Equipment (0-2)** | 1.5 | (0.7) | 2 | (1-2) | 0-2 |
| **Total (0-20)** | 12.6 | (3.0) | 13 | (11-15) | 0-20 |
| **Northwick Park Dependency Scales** | | | | | |
| **NPDS scores (Nursing and care interventions)** | | | | | |
| **Basic care needs (0-65)** | 20.7 | (15.2) | 17 | (8-34) | 0-55 |
| **Special nursing needs (0-35)** | 3.7 | (5.1) | 0 | (0-5) | 0-30 |
| **Total NPDS (0-100)** | 24.4 | (18.7) | 19 | (9-39) | 0-79 |
| **Estimated care hours per week** | 40.6 | (20.0) | 40 | (26-60) | 0-79 |
| **NPTDA scores (Therapy interventions)** | | | | | |
| **Total NPTDA (0-100)** | 26.6 | (7.4) | 26 | (22-31) | 2-57 |
| **Estimated total therapy hours per week** | 22.2 | (8.3) | 21 | (17-26) | 2-95 |

*RCS-v8: Rehabilitation Complexity Scale (version 2), RCS-E-v12: Rehabilitation Complexity Scale-Extended v12*
*NPDS: Northwick Park nursing Dependency Scale, NPTDA: Northwick Park Therapy Dependency Assessment*




**Figures 7 (a) and (b): Distributions of total RCS-v8 and total RCS-E scores**

Spearman correlations (rho) were used to examine the relationship between the RCS-E and the RCS-v8, and between the respective components of the RCS-E and the levels of nursing and therapy intervention. *Table 7* shows the correlations between the different scales. As expected there was a very strong relationship between the total RCS-v8 and RCS-E scores (rho 0.95).

There were very strong correlations between the RCS-E C+N subscale and the total NPDS and care/ nursing hours per week (rho 0.87 and 0.84 respectively, both p<0.001); and between the RCS-E Therapy subscale and the NPTDA and total therapy hours per week (rho 0.71 and 0.70, p<0.001). Weaker correlations were seen between the nursing and therapy elements of the RCS-E (rho 0.27), which were proportionate with the weaker relationship between the nursing and therapy interventions as measured by the Northwick Park Dependency Scales. However, in view of the very large number within this study, all correlations reached statistical significance at the level of p<0.001.

**Table 7: Correlations* between the RCS-v8, the RCS-E-v12 and the Northwick Park Dependency Scale ratings (n=2241)**

| | RCS-v8 | | | RCS-E-v12 | | | Dependency Scales | | |
|---|---|---|---|---|---|---|---|---|---|
| | C+N | TD+TI | Total | RCS-E C+N | RCS-E TD+TI | RCS-E Total | NPDS score | Care hours | NPTDA score |
| **RCS-v8** | | | | | | | | | |
| RCS C+N subscale | | | | | | | | | |
| RCS TD+TI subscale | 0.22 | | | | | | | | |
| RCS Total score | **0.89** | 0.54 | | | | | | | |
| **RCS-E** | | | | | | | | | |
| RCS-E C+N subscale | **0.99** | 0.21 | **0.88** | | | | | | |
| RCS-E TD+TI subscale | 0.27 | **0.91** | 0.55 | 0.27 | | | | | |
| RCS-E Total score | **0.83** | 0.59 | 0.95 | 0.83 | 0.66 | | | | |
| **Northwick Park Dependency Scales** | | | | | | | | | |
| NPDS Total score | **0.87** | 0.23 | **0.80** | **0.87** | 0.28 | 0.77 | | | |
| Care/nursing hours per week | **0.84** | 0.20 | **0.75** | **0.84** | 0.25 | 0.73 | **0.92** | | |
| NPTDA Total score | 0.24 | **0.67** | 0.45 | 0.24 | **0.71** | 0.53 | 0.24 | 0.22 | |
| Therapy hours per week | 0.22 | **0.67** | 0.41 | 0.22 | **0.70** | 0.49 | 0.23 | 0.20 | **0.86** |

**All correlations are significant at the 0.001 level (2-tailed).*

**Legend:** *RCS-v8: Rehabilitation Complexity Scale (version 2); RCS-E: Rehabilitation Complexity Scale (Extended); C+N: Care and Nursing subscale; TD+TI: Therapy Disciplines and Therapy Intensity subscale; NPDS: Northwick Park Dependency Scale; NPTDA: Northwick Park Therapy Dependency Assessment.*

To determine whether the RCS-E provides added discrimination at the higher end of the relevant subscale, Mann Whitney tests were used to compare care/nursing per week (as estimated by the NPDS algorithm) between patients scoring 6/7 and 7/7 on the combined RCS-E Care & Nursing (RCS-E C+N) subscale. Similarly, total therapy hours per week were compared between patients scoring 6/8. 7/8 and 8/8 on the combined RCS-E Therapy subscale (i.e. RCS-E TD+TI). To account for multiple tests, the threshold for significance was taken as p<0.01.

*Figure 8* shows the distribution of care/ nursing hours per week within each level of the relevant RCS-E C+N subscale. There was a clear plateau at the upper end, and patients scoring 7/7 (n=103) on the RCS-E C+N subscale did not have significantly higher care and nursing interventions than those scoring 6/7 (n=85) (z= -1.8, two-tailed p = 0.072).

In contrast, *Figure 9* shows the distribution of therapy hours per week within each level of the relevant RCS-E Therapy subscale. Patients scoring 7/8 (n=383) on the RCS-E TD+TI subscale had significantly higher levels of therapy interventions than those scoring 6/8 (n=551) (z = -12.1, two-tailed p<0.001). Similarly, patients scoring 8/8 (n=241) on the RCS-E TD+TI subscale had significantly higher levels of therapy interventions than those scoring 7/8 (n=383) (z = -8.4, two-tailed p <0.001).

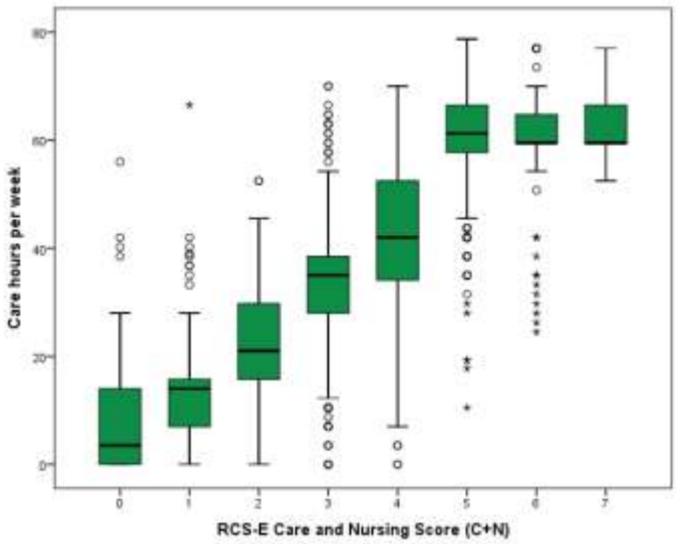



**Figure 8:** **Care hours per week within each level of the summed RCS-E Care and Nursing needs subscales**

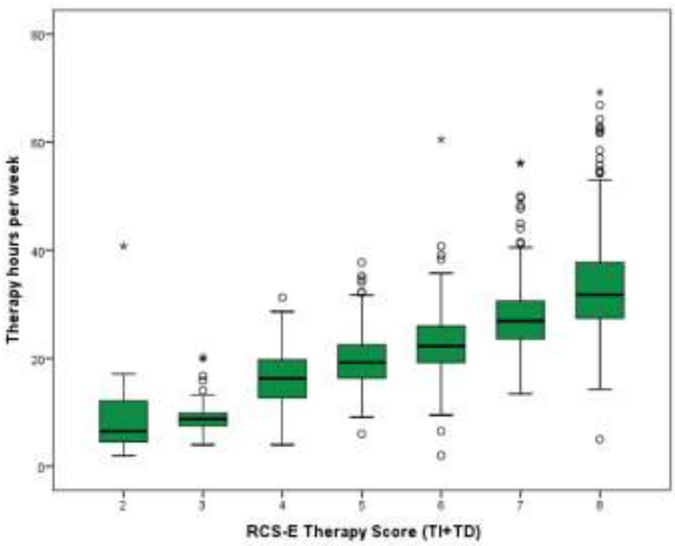



**Figure 9:** **Therapy hours per week within each level of the RCS-E Therapy subscale (Therapy disciplines + intensity).**

*Table 8* shows the results of a principal components factor analysis (PCA) on the correlations of the RCS-E items. Cronbach's alpha for the total scale was 0.64. Item-total correlations ranged from rho 0.46 – 0.74. PCA with varimax rotation indicated a two factor solution accounting for 60% of the variance in total scores.

- The first factor, 'Nursing/medical' care, including equipment, accounted for 37% of the variance. The C, N, M and E items all loaded high (0.58-0.83) on this factor and low (<0.1) on factor 2.
- The second factor, 'Therapy', accounted for 24% of the variance. The two therapy items (TD and TI) both loaded above 0.85 on this factor and low on factor 1.

Cronbach's alphas for these two subscales were 0.69 and 0.67 respectively.

**Table 8: Results of principal components factor analysis with orthogonal rotation on the correlations of the six RCS-E-v12* items using start of treatment scores (n=318)**

| | Un-rotated principal component loading | | Varimax rotation Orthogonal factor loading | |
|---|---|---|---|---|
| **RCS-E item** | **Factor 1** eigenvalue 2.2 | **Factor 2** eigenvalue 1.4 | **Factor 1** | **Factor 2** |
| **Care (C)** | 0.80 | -0.23 | **0.83** | 0.08 |
| **Nursing (N)** | 0.81 | -0.22 | **0.83** | 0.09 |
| **Medical (M)** | 0.54 | -0.26 | **0.60** | -0.05 |
| **Therapy Disciplines (TD)** | 0.29 | 0.83 | -0.03 | **0.88** |
| **Therapy Intensity (TI)** | 0.46 | 0.74 | 0.16 | **0.85** |
| **Equipment** | 0.58 | -0.12 | **0.58** | 0.09 |

**RCS-E: Rehabilitation Complexity Scale-Extended v12*

In summary, there was a very strong correlation between RCS and RCS-E-v12 data, but the RCS-E showed better discrimination of patients with high requirements for therapy intervention above the ceiling of the RCS. However, it did not provide any greater discrimination with respect to needs for staff time in relation to care and nursing, suggesting a ceiling effect still in these elements of the scale.

The factor structure of the RCS-E proved similar to the RCS-v8. That is, it showed moderate internal consistency, suggesting that the six subscales are broadly cumulative. Nevertheless, both exploratory and confirmatory factor analyses suggest that the scale has two distinct dimensions ('Nursing/medical care' and 'Therapy'), the 'Equipment' item being included within the former, rather than the latter. On the other hand, the five components each have differential impact for rehabilitation requirements, and between them provide a profile of rehabilitation needs. Therefore, separate reporting of item scores (e.g. C2 N3 M2 T5 E2) is still recommended to facilitate clinical interpretation.

Containing only one item more than the RCS-v2, the RCS-E still took only 1-2 minutes to rate once the team was familiar with scoring, and was therefore feasible for use in routine clinical practice. Overall, the team preferred the RCS-E as it provided additional clinically important information with respect to therapy and equipment needs, for minimal additional rating burden.

**Limitations of this study** were that, once again, it was conducted on cases from a single Level 1a rehabilitation centre and it did not include any evaluation of the 'risk' scale as the majority of patients required physical care rather than risk management.

We concluded that the RCS-E-v12 offers added value over the RCS in the identification of patients with highly complex therapy and equipment needs, for minimal additional scoring burden, but there were still concerns about ceiling effects in the care/nursing and medical elements of the scale.

## *Equivalence of the RCS and RCS-E*

In 2012 we conducted a parallel analysis to examine the scaling equivalence of the Rehabilitation Complexity Scale-v8 (RCS) and the extended RCS (RCS-E-v12). Parallel complexity ratings using the RCS and RCSE (n=19,307 data pairs) were extracted from the UK Rehabilitation Outcomes Collaborative (UKROC) national database. The data set comprised consecutive patients admitted for in-patient neurorehabilitation between February 2009 and January 2012 to a total of 48 specialist in-patient neurorehabilitation services: (20 Level 1 and 28 Level 2 services). The data were presented as a poster at the American Academy of Physical Medicine and Rehabilitation Annual Assembly.[73]

*Table 9* shows the Median, IQR and Range of RCS-E scores for each scoring level of the RCS-v8. Table 10 shows the final conversion selected to derive RCS bands from the RCS-E v12 scores.

**Table 9: Median, IQR and Range RCS-E-v12 scores for each scoring level of the RCS v8**

| RCS | RCS-E Median | IQR | Range | Final conversion |
|---|---|---|---|---|
| **0** | 0 | 0-1 | 0-2 | **0** |
| **1** | 2 | 1-2 | 1-2 | **1** |
| **2** | 3 | 2-3 | 2-4 | **2** |
| **3** | 4 | 3-4 | 3-5 | **3** |
| **4** | 5 | 4-5 | 4-7 | **4** |
| **5** | 5 | 5-5 | 5-5 | **5** |
| **6** | 7 | 6-7 | 6-9 | **6** |
| **7** | 8 | 7-8 | 7-10 | **7-8** |
| **8** | 9 | 9-9 | 8-11 | **9** |
| **9** | 10 | 10-11 | 9-13 | **10** |
| **10** | 12 | 11-12 | 10-14 | **11-12** |
| **11** | 13 | 12-13 | 11-15 | **13** |
| **12** | 14 | 14-15 | 12-17 | **14** |
| **13** | 16 | 15-16 | 13-18 | **15-16** |
| **14** | 17 | 16-17 | 14-19 | **17** |
| **15** | 18 | 17-19 | 16-20 | **18-20** |

**Table 10: Conversion of RCS groups**

| Group | RCS-v8 | RCS-E-v12 Conversion |
|---|---|---|
| **1** | 0-3 | 0-3 |
| **2** | 4-6 | 4-6 |
| **3** | 7-9 | 7-10 |
| **4** | 10-12 | 11-14 |
| **5** | 13-15 | 15-20 |

## *RCS-E version 13*

To address the remaining ceiling effects, in the next version (Version 13) of the RCS-E we extended the care and medical scales to a range of 0-4, giving the scale structure shown in *Table 3* (see page 32). As the scale structure had not altered significantly we did not re-evaluate the whole scale, but we performed a further exploratory and confirmatory factor analysis to confirm its scale structure.

For this analysis we used the multicentre UKROC cohort collected over 12 months between April 2013-2014. The RCS-E-v13 was recorded in 5014 episodes, mean age 56.3 (sd 18), male:female ratio 58:42; diagnosis: 61% acquired brain injury, 10% spinal cord injury, 11% progressive conditions, 5% peripheral neurological conditions and 11% other conditions. This multi centre cohort included 4881 ratings of the RCS-Risk item, and so addressed the acknowledged weaknesses of the evaluation of RCS-E-v 12.

## Dimensionality

The dataset was randomly split into two groups using SPSS. Exploratory factor analysis was conducted on the first sample and CFA on the second.

The results of EFA for the RCS-E-v13 are shown in *Table 11*. Only one factor had an eigenvalue >1, accounting for 36% of the variance. All six items loaded strongly onto this factor (range 0.51-0.72), and Cronbach's alpha for the full scale was 0.66. A second factor with eigenvalue 0.96 accounted for a further 16% of the variance (52% both together).

With varimax rotation of two factors the scale split neatly into: 'Medical/nursing' and 'therapy' domains. In this extended version, the Equipment item is now included with the therapy dimension.

**Table 11: Results of principal components factor analysis with orthogonal rotation on the correlations of the six RCS-E-v13 items using start of treatment scores (n=5104)**

| | Un-rotated principal component loading | | Varimax rotation Orthogonal factor loading | |
|---|---|---|---|---|
| RCS-E item | Factor 1 eigenvalue 2.2 | Factor 2 eigenvalue 0.96 | Factor 1 | Factor 2 |
| Care (C) | 0.61 | -0.39 | 0.71 | |
| Nursing (N) | 0.63 | | 0.69 | |
| Medical (M) | 0.72 | -0.38 | 0.79 | |
| Therapy Disciplines (TD) | 0.55 | .0.36 | | 0.63 |
| Therapy Intensity (TI) | 0.60 | 0.43 | | 0.72 |
| Equipment | 0.51 | 0.49 | | 0.71 |

*Figure 10:* shows the results of CFA for the two-factor solution. The Goodness of fit index was 0.99, with Comparative fit index (CFI) 0.95 confirming a robust two-dimensional structure.

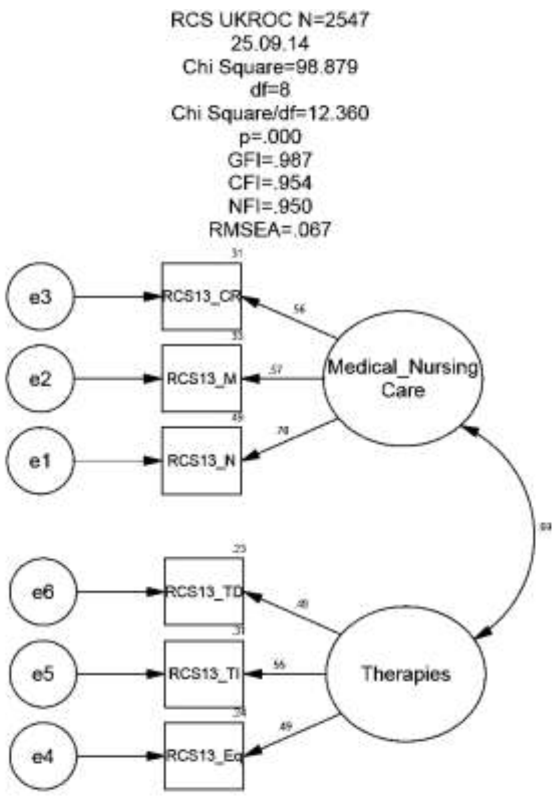



**Figure 10: The results of confirmatory factor analysis for the two-factor solution of the RCS-E-v 13**

**Ceiling effects:**

*Figure 11* shows the distribution of total care and nursing hours per week (as estimated by the NPCNA) within each level of the relevant RCS-E C/R+N subscale. The plateau at the upper end of the scale has now reduced.

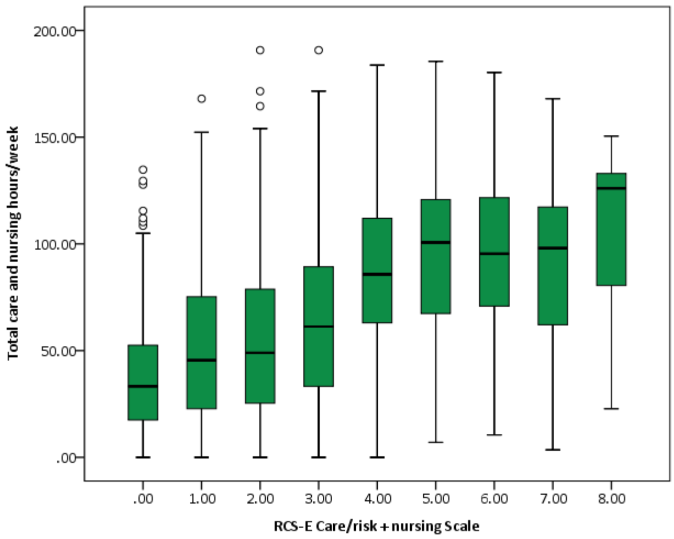



**Figure 11: Care hours per week within each level of the summed RCS-E-v13 care/risk and nursing subscale**

*Figure 12* shows the distribution of therapy hours per week within each level of the relevant RCS-E T subscale. The incremental increase in therapy input hours for each level of the scale is still preserved in this multi-centre sample. We therefore concluded that the expanded RCS-E-v13 had improved sensitivity at the upper end of the scale, in comparison with RCS-E-v12.

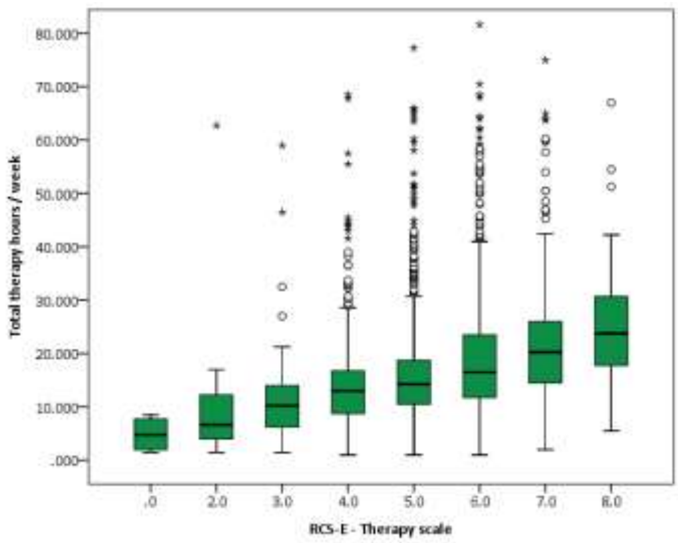



**Figure 12: Therapy hours per week within each level of the summed RCS-E-13 therapy subscales (therapy disciplines+intensity)**

*Figure 13* shows the relationship between the RCS-E-v13 total scores and the total staff (nursing, medical and therapy) hours per week as estimated by the Northwick Park Dependency scales.

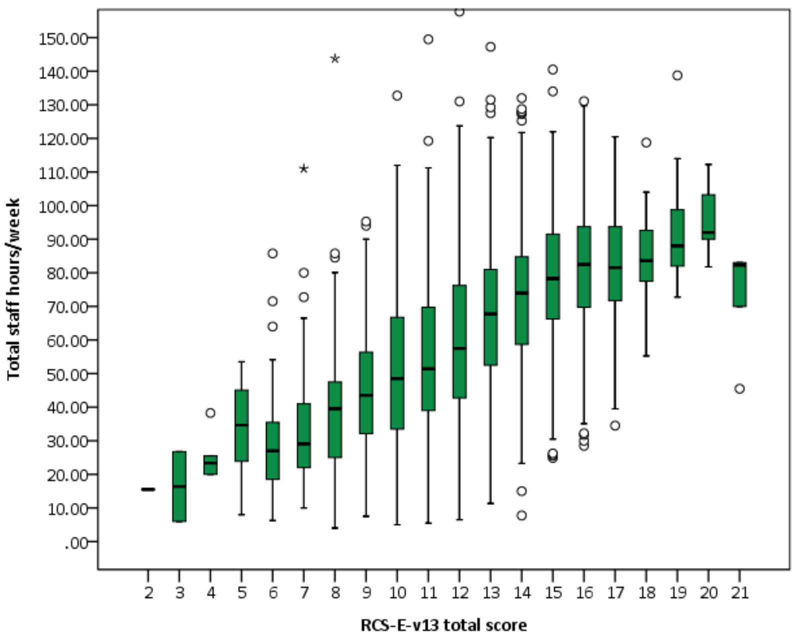



**Figure 13: Distribution of total estimated staff hours per week within each level of the RCS-E-v13 total scores**

## *RCS banding*

The total scores for all three versions of the RCS were highly correlated: RCS-v8 versus RCS-E-v12 rho 0.96 and RCS-E-v12 versus RCS-E-v13 rho=0.99. *Figures 14* and *15* show the relationship between the RCS-v8 and the RCS-E-v13 total scores. As expected there is come expansion towards the upper end of the scale, suggesting again that the RCS-v13 addresses the ceiling effects of the RCS-v 2.

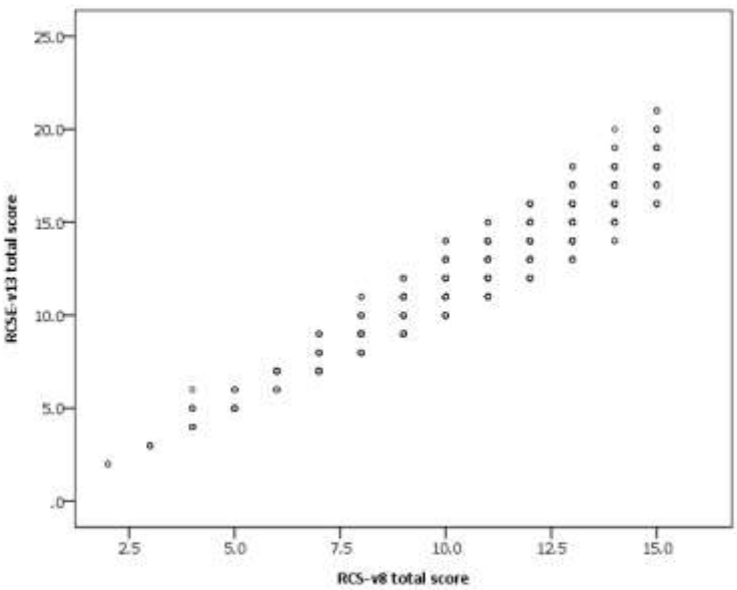



**Figure 14: Scatterplot of the RCS-v8 and the RCS-E-v13 total scores**

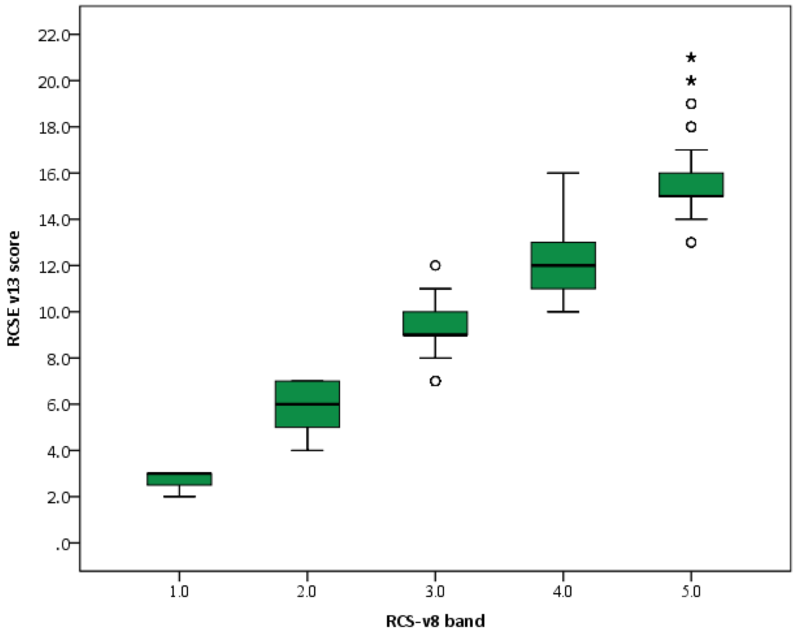



**Figure 15: Distribution of the total RCS-E-v13 scores within each of the complexity bands according to the RCE-v8**

*Table 12* shows the computation of the RCS-E-v13 and RCS-E-v12 scores within the five complexity bands.

**Table 12: RCS-E-v13 and RCS-E-v12 scores within the five complexity bands**

| Band | Very low | Low | Medium | High | Very high |
|---|---|---|---|---|---|
| **RCS-v8** | 0-3 | 4-6 | 7-9 | 10-12 | 13-15 |
| **RCS-E-v12** | 0-3 | 4-6 | 7-10 | 11-14 | 15-20 |
| **RCS-E-v13** | 0-3 | 4-7 | 8-10 | 11-14 | 15-22 |

## *Further development of the RCS-Medical (RCS-M) score*

Development of the RCS and RCS-E up until 2013 focussed on alignment with the nursing and therapy inputs, and added an equipment score to account for the needs for specialist and bespoke equipment. The final piece in the jigsaw is the medical score.

The landscape of Rehabilitation Medicine in the NHS is changing. Due to mounting pressure on hospital beds in the acute services, patients are being transferred to rehabilitation at an earlier stage in their recovery. Inpatient rehabilitation services are thus increasingly required to manage patients with on-going medical and surgical needs. At the same time, the trend towards providing more care in community setting creates a pressure to move in-patient rehabilitation services into non-acute settings, where it may not be possible to care safely for patients who are still medically unstable. It is therefore important for rehabilitation services to be able to describe accurately the medical acuity of their caseload, and the co-dependencies in terms of the facilities and other specialties that their patients may need to access.

The RCS-Medical score had potential application in this context as it provided a simple measure of medical needs– in particular the requirement for specialist investigation and procedures and 24 hour emergency support. However, it lacked the sophistication to provide detailed information for service planning, and this has lead to some further refinement of the RCS-M score.

### *The RCS-ET*

The need for further development of the RCS first arose in the acute trauma pathway. Delayed transfer to rehabilitation is a significant problem for the Major Trauma networks. This may result from waiting lists for rehabilitation services or genuine medical/surgical instability requiring longer lengths of stay in the Major Trauma Centre (MTC). Moreover, some patients may fluctuate over time between instability and fitness for rehabilitation.

The RCS-ET was developed as an extension of the RCS-E to assist in monitoring of delayed transfers for audit purposes. It includes a further expansion of the RCS-M scores to assist with identification the 'R point' – that is, the point in the recovery pathway when the patient is ready to transfer from the acute trauma setting to rehabilitation.

Daily recording of the RCS-ET M-scores may be used to identify the R-Point and monitor continued fitness for transfer. This clinical decision would be made by the Trauma team on the daily ward round:

- At M = 6, the patient still requires care in the specialist MTC setting.
- At M = 5, they require ongoing trauma care, but this can be delivered locally in a Trauma Unit.
- At M = 4, they still have acute medical / surgical needs requiring out of hours care in an acute care setting, or hyper-acute rehabilitation unit.

M1-3 remain as in all other versions of the RCS:

- At M = 3 they have a potentially unstable medical/psychiatric condition – requiring 24 hour availability of on-site acute medical/psychiatric cover.
- At M = 2 , they require specialist medical/psychiatric intervention for diagnosis or management or procedures – requiring in-patient hospital care (in DGH or specialist unit).
- At M = 1, they require basic intervention/monitoring/treatment which could be delivered in a community hospital with day time medical cover.
- At M = 0, no active medical intervention is required – they could be managed by occasional visits by GP.

This refinement of the RCS-E score would be equally applicable in other acute settings (e.g. critical care, neurosciences etc). This further development is sufficiently new that no data have been collected as yet, so that at this point it remains a conceptual design for future evaluation.

Beyond simply recording the level of medical needs, it is also pertinent to record the types of medical inputs required. This information will assist service planners and developers to locate rehabilitation services in settings that provide for the relevant co-dependencies.

# Development of the Medical Activity Assessment (MAA)

The Medical Activity Assessment (MAA) tool is a data collection device designed to complement the basic medical score within the Rehabilitation Complexity Scale (RCSE-M), providing detailed information on the medical resource utilisation of patients undergoing rehabilitation. It was developed in three phases:

- Phase 1 involved the retrospective analysis of medical activity data in a single rehabilitation unit to identify and categorise the types of medical intervention and resources used by patients undergoing inpatient rehabilitation. This analysis conferred content validity and informed the construction of the first version of the MAA.

- During Phase 2, the MAA was prospectively applied in the same unit over a one year period, during which it was iteratively tested and refined.
- Phase 3 involved a brief prospective pilot study in which MAA data were systematically collected over a three–week period in two rehabilitation units to test generalisability. Intra-rater reliability (test-re-test repeatability) was also evaluated in this period.

This development was supported by two academic fellows training in Rehabilitation Medicine, Dr. Anton Pick and Dr. Meenakshi Nayar.

### Phase 1- conducted in the Regional Rehabilitation Unit at Northwick Park Hospital:

In this tertiary Level 1a rehabilitation service, RCS-E data is routinely collected prospectively during weekly medical ward rounds for all current in-patients. Between 2010-3, wherever patients were scored at RCSE-M ≥2, the medical team recorded in free text format the medical activity or event that qualified for the given RCSE-M score. A retrospective thematic analysis of these records was conducted to identify the key themes for medical intervention that would form the main headings of the MAA. The main themes in the activity list were:

- Clinical issues addressed
- Investigations completed
- Procedures
- Specialist Input
- Non-clinical Activity

From this list, the first draft of the MAA was developed to provide a more comprehensive checklist of resources used.

### Phase 2:

During phase 2, the checklist was recorded prospectively, and refined through a process of iterative review.

In Phase 3, preliminary data were collected to quantify and describe medical resource utilisation and to examine the relationship between RCSE-M scores and resource utilisation including:

- Medical time spent during normal working hours,
- On-call alert, when the on-call medical / surgical teams in the hospital were informed of a patient who is medically unstable and may need intervention,
- Actual out-of-hours intervention by the emergency teams.

The checklist was completed at the same time as the RCSE-M score during the weekly medical ward round. Physician time was also recorded using the NPTDA medical score. Test-re-test repeatability was evaluated by recording the checklist in a subgroup of patients (n=43) on two separate occasions at least 4 hours apart. Agreement between the first and second rating was tested using linear-weighted kappa coefficients (kw). During the three-week period, a total of n=87 ratings were collected from a total of 28 patients in two centres.

- The mean age of the sample was 44 years (SD 15), M:F ratio 3:1, mean length of stay 97 days.
- Diagnosis: 25 (89%) acquired brain injury (10 strokes, 23 TBI, 2 hypoxic/other); 1(4%) spinal cord injury; 1(4%) peripheral neurological condition;1(4%) other.
- Intra-rater reliability of the RCS-E-M scores (n=43) showed 100% agreement (kw = 1.0).
- Intra-rater reliability for physician time (n=43) was kw=0.940.

Within this pilot just 3 (4%) of RCSE-M scores were M1;  63 (72%) were M2 and 21 (24%) of scores were M3-4.  The commonest medical issues in the M3-4 group were sepsis and autonomic dysreflexia (‘storming’), occurring in >50% of patients.

- 59% of all investigations were carried out in this subgroup
- 76% of them required input from at least one other specialty
- They were attended to out of hours around ten times more frequently than those patients scoring M 1 and M2.

The most common specialist inputs were from the tracheostomy team (ENT) (15%), neurology (14%), acute medicine (10%), ophthalmology (8%), urology (8%) and orthopaedics (6%), suggesting that these would be important co-dependencies for this group.  Other inputs comprised radiology, neurosurgery, endocrinology, gastroenterology, vascular surgery, microbiology, dermatology, haematology, respiratory and tissue viability services.  One patient with RCS-M=4 required input from eight different specialists during the 2-week period.

In summary, this small pilot study provided evidence for the validity of the RCS-M score as a measure of clinical resource use in in-patient neuro rehabilitation settings.  The MAA proved feasible to collect and provided useful additional information about the types of resources used to facilitate service planning.

# The RCS for specialist nursing homes and slow-stream rehabilitation

Some patients with profound complex neurological disability will require longer periods of rehabilitation in a ‘slow-stream’ rehabilitation facility over a period of 18 months or more.  Others will require life-long care in a specialist nursing home, but nevertheless require a maintenance therapy programme to prevent the progressive development of contractures and deformity that frequently accompany immobility and spasticity if not managed appropriately.  It is as important to be able to measure complexity of rehabilitation needs in this group, but the descriptors are somewhat different.

In 2013, the British Society of Rehabilitation Medicine (BSRM) published guidelines for specialist nursing home care for people with complex neurological disability.[74]  As part of their development, the UKROC programme supported development of an adapted version of the RCS for use in specialist nursing homes (RCS-SNH).

The RCS-SNH has the same six items as other versions of the Rehabilitation Complexity Scale, but has a score range of 0-18 (C 0-3, N 0-3, M 0-3,TD 0-3, TI 0-3, E 0-3).  After a pilot application in two specialist nursing home settings, typical score ranges were used to describe four different categories of needs for care, rehabilitation and equipment/ facilities, and linked to approximate cost-ranges.  Further work is required, however, to evaluate the RCS-SNH in this context.

### Conclusion

Through this iterative development process the Rehabilitation Complexity Scale has developed from a four-item scale (range 0-15) to a six-item scale (range 0-22) with an increasingly linear relationship to resources use measured in terms of staff time, which constitutes the main indicator of cost within a rehabilitation service.

Whilst further work is on-going to test the latest refinements of the medical subscale, the tool is shown to be a robust measure of resource use which reflects the additional costs of treating patients with complex needs and for quantifying differences in caseload complexity between different populations and services.

It therefore has potential use as a case-mix measure for specialist rehabilitation. For current versions of the various Rehabilitation Complexity Scales and associated tools see:
*Appendix 4.1:* the RCS-2 (version 8), *Appendix 4.2:* the RCS-E (version 12), *Appendix 4.3:* The RCS-E (version 13), *Appendix 4.4:* The RCS-SNH, *Appendix 4.5:*The Medical Activity Assessment.

# Categorisation of needs for rehabilitation – the Patient Categorisation Tool

As noted in Chapter 2, the DoH's 2009 Specialised Service National Definition Set (SSNDS)[15 75] identified three levels of service and four categories of patient need. The NHS England Service specification for 'Specialised Rehabilitation for patients with highly complex needs', encompasses patients with category A needs undergoing rehabilitation in Level1/2a services (see *Appendix 2.1*). Whilst the Rehabilitation Complexity Scale provides an overview of a patient's resource requirements, it does not describe the types of need that would support the identification of patients with 'Category A' needs. The Patient Categorisation Tool (PCAT) was therefore developed to fulfil this purpose.

The initial tool was in the form of a simple checklist to identify patients with category A needs based on the descriptors, see *Table 1 of Appendix 2.1*. It was subsequently expanded to an ordinal scale, in order to distinguish patients with category A, B or C needs.

The PCAT tool comprises 16 items, each rated on a score of 1-3, and 1 item (duration) rated 1-2, giving a total score range of 17-50, (see *Appendix 4.6*). It was first introduced in 2012, and the facility for full itemised recording was available within the UKROC database from April 2013.

In the multi-centre cohort analysis presented below we examined the value of the PCAT as a tool to identify patients with category A Needs. Exploratory and confirmatory factor analyses were conducted to determine its factor structure. We also examined inter-rater reliability, its concurrent relationship with other measures, and explored possible cut-off points that may reasonably identify category A activity within the sample.

The analysis was conducted on all PCAT data recorded in the UKROC database for all Level 1 and 2 services between April 2012 and December 2014. During this period, PCAT ratings were recorded on admission for a total of 5396 patients, comprising 63.1% of the cohort. For the purpose of the factor analyses, the sample was randomly divided into two approximately equal sub-samples using the randomisation facility in SPSS v22. The demographics of all three samples are shown in *Table 13.* No significant differences were found between the two subsets, confirming that the randomisation process was successful.

**Table 13: Demographics of the total sample and randomised subsamples**

| | **Total sample N=5396** | **Sub-sample A N=2691** | **Sub-sample B N=2705** |
|---|---|---|---|
| **Mean age (SD) years** | 54.4 (18.2) | 54.5 (18.4) | 54.4 (18.2) |
| **Males : Females %** | 58:42% | 58:42% | 58:42% |
| **Length of stay (SD) days** | 77.9 (67.3) | 77.4 (66.6) | 78.4 (68.1) |
| **Diagnostic categories (%):** | | | |
| **Acquired brain injury** | 66.4% | 65.2% | 67.6% |
| **Spinal cord injury** | 9.4% | 9.4% | 9.3% |
| **Peripheral neurological** | 5.0% | 4.7% | 5.3% |
| **Progressive** | 9.9% | 10.6% | 9.3% |
| **Mean Total PCAT score (SD)** | **29.7 (6.8)** | **29.7 (6.7)** | **29.7 (6.9%)** |

# Dimensionality and internal consistency

**Exploratory factor analysis** was conducted on Sub-sample A, using principal components analysis with Varimax rotation to examine the dimensionality and internal consistency of the PCAT. The findings are summarised in *Table 14*.

Item total correlations ranged from 0.31-0.72. Principal components analysis showed that all but two items loaded strongly into the first principal component with loadings ≥0.45, the exceptions being tracheostomy (0.3) and vocational rehabilitation (0.29). Cronbach's alpha was 0.88 for the total scale.

Inspection of the Scree plot suggested a 2-factor solution. Varimax rotation revealed two factors.

- The first "*Cognitive/psychosocial"* factor (PCAT-Cog) comprised needs relating to psychiatric input, cognitive, behavioural and mood management. Also within this factor were 'Emotional load on staff', and the needs for 'Family support', 'Vocational rehabilitation' and 'Medico-legal' input (such as assessment of mental capacity etc).
- The second *'Complex Physical'* factor (PCAT-Phys) comprised needs for medical input, physical handling, and management of disability, tracheostomy and swallow/nutrition. Also within this factor were needs for special facilities, and more intensive or longer duration rehabilitation.
- Needs for communication and discharge planning loaded strongly on both factors.

**Table 14: Results of exploratory factor analysis**

| Item | Mean (SD) score | | Item total correlation | Loading on 1st principal component | Varimax rotation | |
|---|---|---|---|---|---|---|
| | | | | | Factor 1 | Factor 2 |
| **Medical** | 2.19 | (0.63) | 0.50** | 0.50 | | **0.45** |
| **Physical handling** | 2.40 | (0.68) | 0.47** | 0.45 | | **0.69** |
| **Tracheostomy** | 1.09 | (0.40) | 0.31** | 0.30 | | **0.51** |
| **Swallow/nutrition** | 1.56 | (0.72) | 0.63** | 0.63 | | **0.64** |
| **Disability** | 1.90 | (0.48) | 0.48** | 0.48 | | **0.63** |
| **Facilities** | 1.70 | (0.77) | 0.56** | 0.53 | | **0.71** |
| **Duration** | 0.67 | (0.66) | 0.61** | 0.61 | | **0.58** |
| **Intensity** | 2.26 | (0.78) | 0.66** | 0.64 | (0.40) | **0.53** |
| **Discharge planning** | 2.10 | (0.66) | 0.64** | 0.65 | (0.46) | **0.47** |
| **Communication** | 1.80 | (0.85) | 0.70** | 0.69 | **0.49** | (0.49) |
| **Psychiatric** | 1.42 | (0.63) | 0.42** | 0.48 | **0.73** | |
| **Cognitive** | 2.10 | (0.84) | 0.65** | 0.64 | **0.63** | |
| **Behaviour** | 1.48 | (0.66) | 0.56** | 0.59 | **0.81** | |
| **Mood** | 1.72 | (0.69) | 0.58** | 0.59 | **0.70** | |
| **Family support** | 1.94 | (0.60) | 0.64** | 0.65 | **0.54** | |
| **Staff emotional** | 1.62 | (0.71) | 0.72** | 0.74 | **0.73** | |
| **Vocational rehab** | 1.61 | (0.79) | 0.35** | 0.29 | **0.34** | |
| **Medico legal** | 1.53 | (0.74) | 0.63** | 0.63 | **0.58** | |
| **Cronbach's alpha** | | | | | **0.83** | **0.79** |

*** All significant at p<0.001*

### Confirmatory factor analysis

To determine the reliability of the hypothesised two-factor model yielded by exploratory factor analysis, subsample B (N=2705) was examined using confirmatory factor analysis. The model was specified to estimate each of the loadings on the two-factor hypothesised model (Table 14), with 'Communication' in factor 1 (PCAT-Cog) and 'discharge planning' in factor 2 (PCAT-Phys). Inspection of the modification indices suggested that model fit was significantly improved if 'Discharge planning' and 'Communication' were allowed to load on both the factors. For the final model the Root Mean Square Error of Approximation (RMSEA) was 0.087, Comparative Fit Index/Tucker/Lewis Index (CFI/TLI) 0.831/0.807 and the Goodness of fit Index (GFI) was 0.882.

A slight improvement in fit of the model could have been achieved by adding covariances between error terms of three pairs of items ('Communication' and 'Cognitive'; 'Behavioural' and 'Psychiatric'; 'Facilities ' and 'Physical handling'). This model produced the following fit statistics RMSEA was 0.073, CFI/TLI 0.883/0.863 and the GFI was 0.919. The final model therefore supported the two-factor hypothesised structure of the PCAT scale, but as they were not substantially different we decided to retain the simpler model, as illustrated in *Figure 16*.

In summary, factor analysis suggests that the PCAT may reasonably be summed into a single total score, but also comprises two factors – one relating principally to cognitive / psychosocial requirements (PCAT-Cog) and the other to physical requirements (PCAT-Phys). It is possible that the PCAT-Cog subscale may give a better indication of needs for specialist cognitive/behavioural rehabilitation services, where patients are likely to have fewer needs for physical rehabilitation.

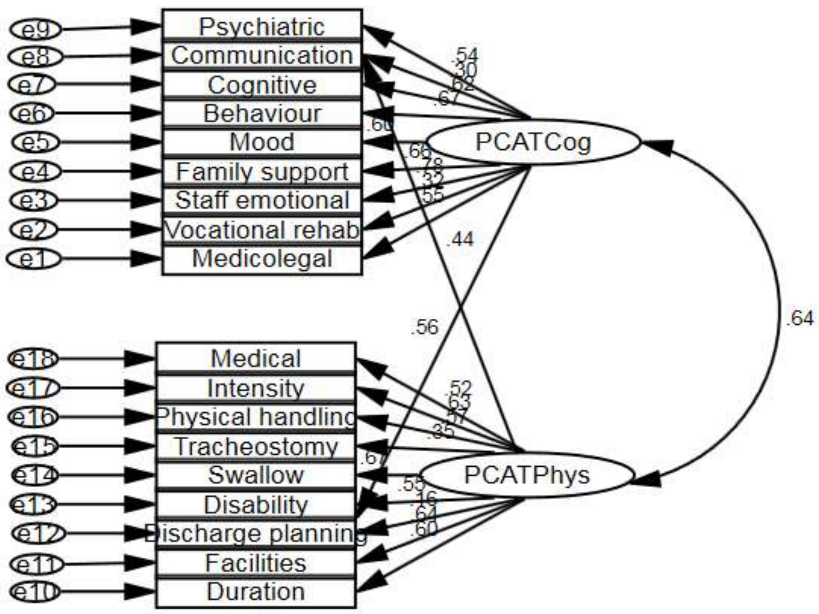



**Figure 16: Confirmatory factor analysis. Model specification.**

## Reproducibility

On the Regional Rehabilitation Unit at Northwick Park, the medical team routinely undertake PCAT ratings on admission, based on presentation of findings collected at the admission medical clerking within 48 hours of admission for n=41 patients. To test inter-rater reliability, two consultants independently recorded PCAT scores for a sub-sample of n=41 patients. No systematic rating basis was observed between the raters – both of whom rated a median total of 35 per patient.

*Table 15* shows the level of agreement between their item-level ratings expressed as the % absolute agreement, and Cohen's kappa coefficients. Kappa values were unweighted as the PCAT items are scored on only three levels. The percentage absolute agreement ranged from 66-90% and Kappa coefficients ranged from 0.50-0.80. Thus,16 items showed 'moderate–substantial' agreement and one (tracheostomy) showed 'almost perfect' agreement. The Intra-class correlation coefficient (ICC) for the total PCAT scores was 0.96 (95% CI 0.93, 0.98).

**Table 15: Item level agreement between two raters**

| PCAT item | Kappa value | Std. Error | 95% CI Lower | 95% CI Upper |
|---|---|---|---|---|
| Medical/neuropsych needs | 0.64 | 0.11 | 0.43 | 0.84 |
| Therapy intensity | 0.61 | 0.10 | 0.41 | 0.82 |
| Physical | 0.80 | 0.09 | 0.63 | 0.96 |
| Tracheostomy | 0.86 | 0.12 | 0.63 | 1.00 |
| Swallowing/Nutrition | 0.79 | 0.09 | 0.62 | 0.96 |
| Communication | 0.80 | 0.08 | 0.64 | 0.96 |
| Cognitive | 0.55 | 0.12 | 0.32 | 0.78 |
| Behavioural | 0.71 | 0.14 | 0.44 | 0.98 |
| Mood | 0.50 | 0.12 | 0.28 | 0.73 |
| Disability management | 0.76 | 0.11 | 0.54 | 0.98 |
| Social discharge | 0.58 | 0.12 | 0.35 | 0.81 |
| Family support | 0.70 | 0.10 | 0.50 | 0.90 |
| Emotional load on staff | 0.59 | 0.12 | 0.36 | 0.81 |
| Vocational Rehabilitation | 0.69 | 0.12 | 0.45 | 0.93 |
| Medico-Legal | 0.50 | 0.12 | 0.27 | 0.73 |
| Equipment/Facilities | 0.52 | 0.11 | 0.30 | 0.74 |
| Anticipated duration of stay | 0.52 | 0.14 | 0.25 | 0.79 |

## Validity

Concurrent validity was explored through the relationship of the PCAT tool with measures of complexity (RCS-E-v12), dependency (NPDS and UK FIM+FAM). We did not expect to find a very close relationship as these tools measure different parameters. Nevertheless we expected and found a congruent relationship (positive correlation) with complexity (total RCS score) and dependency (total NPDS score and a discriminant relationship (negative correlation) with independence (Total UK FIM+FAM).

We also expected and found stronger negative relationships between the PCAT Phys and the FIM+FAM Motor score, and between the PCAT-Cog and FIM+FAM Cognitive score, than with their respective counterparts. *Table 16* summarises the Spearman rho correlations between these different parameters.

**Table 16: Spearman rho correlations between PCAT, RCS-E-v12, NPDS and FIM+FAM scores**

| | PCAT-Total | PCAT-Phys | PCAT-Cog |
|---|---|---|---|
| RCS-E-v12 total score (n=5047) | 0.59 | | |
| NPDS Total Score (n=4442) | 0.49 | | |
| FIM+FAM total score (n=4920) | -0.56 | | |
| FIM+FAM motor score (n=4920) | | -0.62 | -0.22 |
| FIM+FAM cognitive score (n=4920) | | -0.43 | -0.55 |

*Because of the large numbers, all correlations were significant at p<0.0001*

## Sensitivity and specificity

No gold standard exists for determining the category of rehabilitation need. Until now, this has been determined by subjective impression on the part of the assessor, with no structured frame of reference. It was nevertheless relevant to examine the relationship between this more structured assessment and the assessor's subjective view.

In addition to rating the PCAT score, assessors were also asked to give their clinical impression of whether the patient had category A, B or C needs. *Figure* 17 shows the distribution of total and sub-scale PCAT scores in relation to this subjective categorisation, and *Table* 17 summarises the descriptive statistics.

These findings suggest that PCAT total scores of 24 and 30 respectively mark the cut-off points for category B and A needs. Equivalent cut-off points for the PCAT-Phys would be 13 and 16, and for the PCAT-Cog would be 11 and 15.

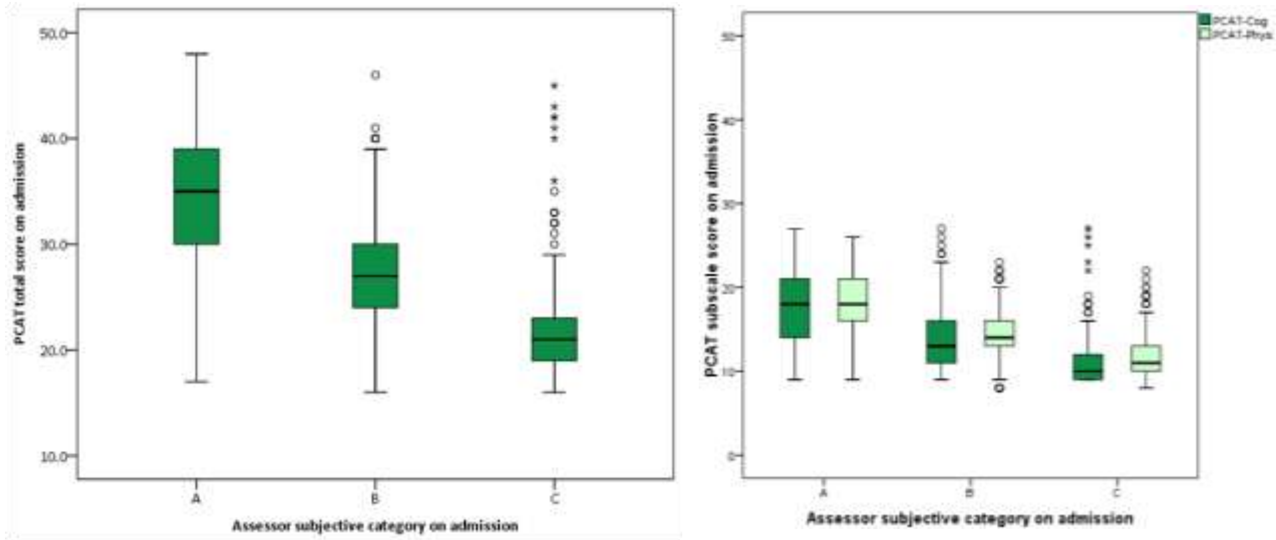
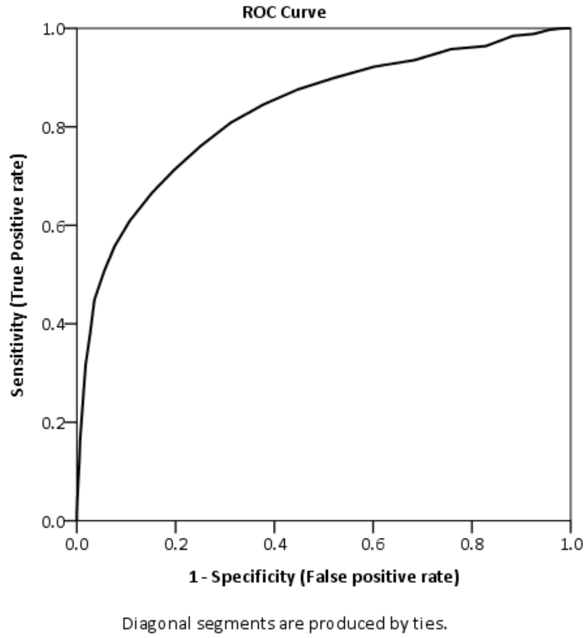



**Figure 17:** **Distribution of PCAT scores in relation to the assessors' subjective categorisation**

**Table 17:** **Descriptive statistics of PCAT scores by category needs**

| | N | Mean | SD | Median | IQR | Min | Max |
|---|---|---|---|---|---|---|---|
| **PCAT total** | | | | | | | |
| **A** | 2392 | 34.1 | (6.3) | 35 | 30-39 | 17 | 48 |
| **B** | 2401 | 27.0 | (4.4) | 27 | 24-30 | 16 | 46 |
| **C** | 413 | 21.5 | (4.2) | 21 | 19-23 | 16 | 45 |
| **PCAT-Phys** | | | | | | | |
| **A** | 2392 | 18.1 | (3.5) | 18 | 16-21 | 9 | 26 |
| **B** | 2401 | 14.5 | (2.6) | 14 | 13-16 | 8 | 23 |
| **C** | 413 | 11.6 | (2.4) | 11 | 10-13 | 8 | 22 |
| **PCAT-Cog** | | | | | | | |
| **A** | 2392 | 17.5 | (4.4) | 18 | 14-21 | 9 | 27 |
| **B** | 2401 | 13.7 | (3.0) | 13 | 11-16 | 9 | 27 |
| **C** | 413 | 11.0 | (2.6) | 10 | 9-12 | 9 | 27 |

*Figure 18* shows the ROC curve for PCAT total score on admission with Assessor subjective rating Category A needs as the state variable. The area under the curve was -0.83 suggesting that the total PCAT on admission is a good test for identifying patients with category A needs. Nevertheless the PCAT-Cog may potentially have some advantage for identifying category A needs within the cognitive behavioural (Level 1c) services where physical needs are expected to be lower.




| Cut-point | True -ves | False +ves |
|---|---|---|
| 28 | 0.85 | 0.38 |
| 29 | 0.81 | 0.31 |
| 30 | 0.76 | 0.25 |
| 31 | 0.71 | 0.20 |
| **Area under the curve: -0.83** | | |

**Figure 18: ROC curve for PCAT total score with Assessor subjective rating Category A needs as the state variable.**

The areas under the ROC curve for PCAT-Phys and PCAT-Cog were 0.78 and 0.82 respectively, suggesting that the total PCAT score provides a more sensitive and specific measure of category A needs (according to subjective evaluation by the assessors) than either of its subscales alone. A PCAT score of ≥30 identified category A patients with a sensitivity of 73% and specificity of 75%; a positive predictive value of 76% and a negative predictive value of 72%.

## Overall summary of clinimetric properties

An overall summary of the clinimetric properties of the PCAT according to the Medical Outcomes Trust Framework is given in Table 18.

**Table 18: Summary of clinimetric properties of the Patient Categorisation Tool (PCAT)**

| Attribute and criteria | Evaluation |
|---|---|
| **Conceptual and measurement model:** ***The rationale for and description of the concept and the populations that the measure is intended to assess*** | |
| **Clinical content and design** | The PCAT is an 18-item ordinal scale, scored on a range of 0-50.<br>Designed to provide a description of an individual's needs for specialist rehabilitation and to distinguish patients with category A, B and C needs who are likely to require respectively Levels 1, 2 and 3 services. |
| **Dimensionality** | Exploratory and confirmatory factor analyses demonstrate that the PCAT can reasonably be summed to a single total score, but that it also has two distinct dimensions (PCAT-Phys) and (PCAT-Cog) of 9 items each). |
| **Reliability:** ***The degree to which the instrument is free from random error*** | |
| **Internal consistency** | Cronbach's alpha 0.88 and item-total correlations (0.31-0.72) showed moderate internal consistency of the whole scale (PCAT-Phys $\alpha$=0.79; PCAT-Cog $\alpha$ = 0.83). |
| **Reproducibility** | Inter-rater agreement between two consultants: ICC for the total PCAT scores 0.96. Item-by item analysis showed absolute agreement ranging from 66-90%. Unweighted Kappa coefficients ranged from 0.50-0.80, representing moderate-substantial agreement. |
| **Validity:** ***The degree to which the instrument measures what it purports to measure*** | |
| **Content** | Based on the criteria set out in the DoH's national definition set for specialised rehabilitation services 2009. |
| **Criterion-related** | No accepted gold standard currently exists. Sensitivity analysis was performed in relation to subjective expert clinical assessment of rehabilitation needs.<br>The area under the ROC curve was -.833 for the total scale.<br>A PCAT score of ≥30 identified category A patients with a sensitivity of 73% and specificity of 75%; a positive predictive value of 76% and a negative predictive value of 72%. |
| **Concurrent** | Convergent and discriminative validity: Strongly positive correlations were seen with measures of complexity and dependency (0.49-0.59, p<0.001) and negative correlation with measures of functional independence (rho -0.56, p<0.001) |
| **Responsiveness:** ***Ability to detect change over time where real changes occur*** | |
| **Change: admission to discharge** | Not tested in this analysis |
| **Interpretability:** ***The degree to which easily understood meaning can be assigned to quantitative scores*** | |
| **Clinical meaning** | Cut-off points for category A and B needs on the total PCAT score were 24 and 30 respectively.<br>Equivalent cut-off points on the sub-scale scores were:<br>PCAT-Phys 13 and 16<br>PCAT-Cog 11 and 15 |
| **Burden:** ***The time, effort or other demands of administering the instrument*** | |
| **Time to administer** | The PCAT tool took <3 minutes to administer by clinicians who used the score regularly in routine practice. |
| **Alternative modes of administration** | None currently available. |
| **Cultural and language adaptations** | None currently available. |

**In conclusion,** this evaluation demonstrates that the PCAT is a reasonably valid and reliable tool for describing complex needs for rehabilitation and identifying those patients with category A or B needs who are likely to require the higher-level facilities and skills of a specialist rehabilitation service.

It is not expected that a PCAT of ≥30 would identify all such patients, as an individual with complex needs in just one or two domains could require specialist facilities but still rate <30 on the total score. Nevertheless, it provides a potentially useful rule of thumb and a basis on which to compare the needs profile of different patient populations.

# Measuring inputs - The Northwick Park Dependency Scales

## *Overview*

The Northwick Park nursing Dependency and Care Needs Assessment (NPDS/NPCNA)[49, 76] and Therapy Dependency Assessment (NPTDA)[53] are tools to quantify nursing and therapy inputs in a rehabilitation setting. They translate directly into an estimated requirement for nursing and therapy time. In addition, they provide more detailed evaluation, not only of the amount of staff time involved, but also information about the types of input and interventions delivered.

## *The Northwick Park Nursing Dependency Scale and Care Needs Assessment (NPDS/NPCNA)*

The Northwick Park Dependency Scale (NPDS) is a rating scale developed to quantify an individual's needs for nursing care and support, particularly in highly dependent patients. It is an ordinal measure rated on a scale of 0-100, divided into two sections:

- **The Basic Care Needs (BCN)** section (range 0-65) includes a total of 16 items associated with activities of daily living such as washing, dressing, eating and drinking - also safety awareness, behaviour and communication. Each item is rated on a Likert scale of 0-3, 0-4 or 0-5.
- **The Special Nursing Needs (SNN)** section (range 0-35) contains seven care items which would normally need to be undertaken by a qualified nurse, or a specially trained carer. These are scored on a dichotomous scale of 0 or 5 to reflect the intensity of nursing input that they represent.
- **Nursing/care hours:** The UKROC software applies the algorithm to derive the Northwick Park Care Needs Assessment (NPCNA)[51] which includes:
    - A daily timetable of when care needs arise,
    - an estimation of care hours per week,
    - the approximate weekly cost of the care package that would be required to meet their needs for basic care and nursing in the community

First published in 1998, the NPDS is increasingly widely used in the UK[57] and has also been trialled in other countries.[57, 77-81] It has been shown to correlate well with other measures of dependency including the Barthel Index (BI)[79, 80, 82] and the Functional Independence Measure (FIM™)[78] but it also provides additional information about the needs for nursing care in clinical rehabilitation settings:[79]

It provides a direct assessment of the number of carers and time taken to complete care tasks, and is designed to be sensitive for highly dependent patients who fall beneath the floor of the BI and FIM.

- It addresses the need for input from qualified nurses as well as basic care.

- It includes assessment of needs for help with cognitive functions such as communication, behavioural management and safety awareness, which often occupy a significant proportion of nursing time in neurorehabilitation.
- It translates, by way of a validated algorithm, into a) a timetable of care needs [51] which may be used to directly plan care packages on discharge from hospital and b) an estimate of care hours and costs which has been used to demonstrate the cost-efficiency of rehabilitation for highly dependent patients.[32] For the NPDS, along with an example of a score sheet, timetable and summary of care hours and costs, see *Appendix 4.7*.

Prior to the start of this programme, the NPDS had already been in use for over a decade. However, it was pertinent to examine what is known about its psychometric properties, and to consider what, if any, further developments are required to maximise its usefulness as a clinical measure.[50] As part of this programme we:

- Conducted a brief systematic review of the existing literature on its psychometric performance and summarise these studies with respect to the Medical Outcomes Trust criteria for a psychometrically robust tool.
- Undertook further analyses in a large dataset gathered prospectively from a cohort of neurorehabilitation inpatients. We evaluated several important psychometric attributes not previously examined, and also explored the relationship between the NPDS, the BI, FIM™ and the UK FIM+FAM[83] to examine the interaction of physical and cognitive elements of the scale.

The paper describing this evaluation has been published in the Journal of Rehabilitation Medicine[50] so the findings will be summarised briefly here.

## *Systematic Review of NPDS Psychometric Studies*

### Methods

To identify existing literature on psychometric evaluation of the NPDS we searched the following databases using the term 'Northwick Park Dependency.mp': Medline and PubMed 1995 – May 2009, Embase 1980-2009, British Nursing Index and Archive 1995-2009, Allied and Complimentary Medicine (AMED) 1995-2009.

### Results

Of a total of 16 articles recovered, five that specifically examined psychometric properties of the NPDS[49, 78-80, 82] are summarised in *Table 19*. These studies largely focused on inter- and intra-rater reliability and concurrent validity. They provide good support for the reliability and concurrent validity of the NPDS, although reliability for the SNN scale could be low, reflecting the dichotomous scoring structure and also the quite specialised nursing needs it captures.

Three other papers[81, 84, 85] recorded the NPDS in conjunction with other parameters of nursing intervention (e.g. care needs, observed care and nursing activities, work sampling) and provide general confirmation of its content and relevance. However, we were unable to identify any published articles addressing internal consistency, dimensionality/factor structure, responsiveness to change or discriminatory power.

**Table 19: Summary of published literature on psychometric properties of the Northwick Park Dependency Scale (NPDS)**

| Author/year | Attributes examined | Sample | Main findings |
|---|---|---|---|
| **Turner-Stokes et al 1998 (source ref) (UK)** | Validity (Barthel Index) and Inter- and intra-rater reliability | In-patient neuro-rehabilitation sample. Three nurses rated 21 patients on three occasions ( 63 paired ratings) | **Inter rater reliability:**<br>Total scores (Spearman correlation): BCN 0.91, SNN 0.68, Total NPDS 0.90 ($p<0.01$)<br>Item by item analysis: (BCN items only)<br>Absolute agreement 38.7 - 92.1%; Agreement ±1 79.0-100%<br>**Intra-rater reliability:**<br>Total scores (Spearman correlation): BCN 0.95, SNN 0.81, Total NPDS 0.93 ($p<0.01$)<br>Item by item analysis: (BCN items only)<br>Absolute agreement 61.9 – 93.0%; Agreement ±1 88.4 - 100%<br>**Validity:**<br>Barthel Index: rho -0.91 ($p<0.01$) |
| **Post et al 2002 (the Netherlands)** | Construct validity (Cronbach's alpha)<br>Validity (Barthel Index)<br>Sensitivity to change<br>Nurses preference | A prospective longitudinal study of stroke patients (n=31) with serial measurements one month apart | **Validity:**<br>Construct validity (Cronbach's alpha): BCN score 0.85-0.92<br>Barthel Index: BCN rho -0.85 to -0.95 at each measurement; SNN rho -0.28 to -0.57; total NPDS score -0.87 to -0.95<br>Strong relationship to global rating of nursing dependency (rho -0.82) and sensitivity to change Wilcoxon z 4.06 ($p<0.001$)<br>Nurses preference: NPDS was not significantly superior to the Barthel Index but 9/12 nurses preferred it for future use |
| **Hatfield et al 2003 (UK)** | Validity (Barthel Index and recorded nursing hours)<br>Inter- rater reliability<br>Utility: time to score | In-patient neuro-rehabilitation sample (n=22). Two raters ( a doctor and a nurse) | **Inter rater reliability:**<br>Total scores (Spearman correlation): BCN 0.92, SNN 0.48, Total NPDS 0.92 ($p<0.01$)<br>Item by item analysis:<br>BCN items: Absolute agreement 55-95%; Agreement ±1 82 - 100%<br>SNN items: Absolute agreement 55-100%.<br>**Validity:**<br>Barthel Index: BCN rho -0.95 ($p<0.01$); Total NPDS rho -0.89 ($p<0.01$);<br>Nursing time: BCN rho -0.88 ($p<0.01$); Total NPDS rho -0.87 ($p<0.01$);<br>Mean time to score NPDS: 4-5 minutes. |
| **Svensson et al 2005 (Sweden)** | Validity (FIM) and Inter- and intra-rater reliability | Patients with brain injury (n=40) in three rehabilitation centres, rated by two nurses (n=13) or and a nurse and an occupational therapist (O/T) (n=27) | **Inter rater reliability:**<br>Item by item analysis: (BCN items only)<br>Nurse-Nurse: Absolute agreement 77 - 100%; unweighted kappa 0.63-1.0<br>Nurse-O/T: Absolute agreement 54-96%; unweighted kappa 0.28-0.80<br>**Intra-rater reliability:**<br>Item by item analysis: (BCN items only)<br>Absolute agreement 71-100%; unweighted kappa 0.53-1.0<br>**Validity:**<br>FIM: Goodman Kruskals's gamma - Nurses: -0.83 (ase* 0.04); O/T : -0.87 (ase 0.04); |
| **Platinga et al 2006 (the Netherlands)** | Validity (Barthel Index and Dutch Care Dependency Score (CDS)) | Mixed rehabilitation population (total n=154) | **Validity:**<br>Total Barthel Index and total NPDS: Mean group correlation (rho) -0.87; percentage explained variance $R^2 = 0.76$. (Within each of the disease groups correlation varied from -0.93 ($R^2 = 0.86$) to -0.70 ($R^2 = 0.49$) which exceeded their criterion of rho 0.60). Total CDS and total NPDS score: rho-0.74 ($R^2 = 0.55$) |

*Ase = asymptotic standard error; BCN = Basic Care Needs subscale for the NPDS; SNN = Special Nursing Needs subscale of the NPDS; FIM= Functional Independence Measure*

## *Psychometric Evaluation of the NPDS in a large neurorehabilitation cohort*

### Participants

Participants were 569 consecutive patients admitted to a single tertiary (Level 1a) specialist post-acute neurorehabilitation unit over a 10 year period between 1999 and 2008. The mean age of the sample was 44.4 year (SD 14.3), M:F ratio approximately 60:40, mean length of stay 99 days, (SD 61) and mean time since injury 154 days (SD 444). Diagnoses were 83% acquired brain injury (of which 68% had stroke, 20% were traumatic, and 12% had other aetiologies), 9% had spinal cord injury and 5% had a peripheral neurological disorder.

### Methods

Routinely rated admission and discharge scores for the NPDS and FIM+FAM were extracted. The UKROC software automatically derives a Barthel Index[86] and FIM score from the FIM+FAM ratings Patients are routinely divided into three dependency groups, based on their admission NPDS scores.[32]

- Low dependency (NPDS score 0-9) patients are largely independent for self-care or require incidental help only.
- Medium dependency (NPDS 10-24) generally require help from one person for most tasks.
- High dependency (NPDS 25-40) generally require help from two or more people for most tasks.
- Very high dependency (NPDS >40) profoundly disabled patients with very complex care needs.

Psychometric analysis was conducted in line with the methodology described earlier in this chapter.

Although the NPDS provides long ordinal data (range 0-100) with subscales ranging from 0-65 (Basic care needs (BCN) scale) and 0-35 (Special nursing needs (SNN) scale, the data are typically skewed. Arguably, therefore non-parametric statistics should be used for analysis of responsiveness and construct validity. In this evaluation we also explored whether parametric or non-parametric analysis made a difference to the conclusion.

- For the evaluation of responsiveness we recorded both effect size (parametric) and Wilcoxon z scores (non-parametric).
- For construct validity we recorded both Pearson r (parametric) and Spearman rho (non-parametric) correlation tests.

### Results

#### *Reliability and Item-Total Correlations*

For the full 23-item NPDS scale coefficient alpha was 0.90, which fell within the desirable range of 0.7-0.95.[87] Coefficient alpha for the BCN scale was 0.93, but only 0.50 for the 7-item SNN section.

Item-total correlations for the BCN scale are presented in *Table 20*. All BCN item-total correlations were above 0.30 and 75% were above 0.50. Item-total correlations for the dichotomous SNN scale were substantially lower, ranging from 0.01 - 0.40. However, this is likely to reflect the very small number of 'Yes' responses for some SNN items as follows: Tracheostomy 23 (4%), Open Wound 55 (10%), Night-time interventions 123 (22%), Psychological support 109 (19%), Isolation 86 (15%), Medical surgical 41 (7%), and Specialing 9 (2%).

**Table 20: Item Means, Item-Total Correlations and Loadings for the first principal component and 3-factor Varimax rotation of NPDS 16 BCN Items***

| ITEM | Range* | Mean | (SD) | Item-Total Correlations | 1st Principal Component | Factor 1 Physical Care 50% variance | Factor 2 Cognitive/Behaviour 12% variance | Factor 3 Eating/Drinking 8% variance |
|---|---|---|---|---|---|---|---|---|
| BASIC CARE NEEDS | | | | | | | | |
| 1. Mobility | 0-4 | 1.70 | (1.34) | 0.77 | 0.82 | **0.77** | | |
| 2. Bed Transfers | 0-3 | 1.17 | (1.15) | 0.80 | 0.84 | **0.85** | | |
| 3a. Toileting assistance: bladder | 0-4 | 1.21 | (1.15) | 0.54 | 0.60 | **0.54** | | (0.51) |
| 3b. Urinary incontinence | 0-3 | 0.50 | (0.89) | 0.61 | 0.66 | | | |
| 4a. Toileting assistance: bowels | 0-5 | 2.22 | (1.90) | 0.85 | 0.89 | **0.88** | | |
| 4b. Bowel incontinence | 0-3 | 0.57 | (0.98) | 0.65 | 0.69 | **0.61** | | |
| 5. Wash/ Groom | 0-5 | 1.48 | (1.20) | 0.78 | 0.82 | **0.64** | (0.52) | |
| 6. Bath/ Shower | 0-5 | 2.30 | (1.51) | 0.87 | 0.91 | **0.85** | | |
| 7. Dressing | 0-5 | 2.13 | (1.45) | 0.88 | 0.91 | **0.86** | | |
| 8a. Eating | 0-3 | 0.67 | (0.72) | 0.36 | 0.42 | | | (0.83) |
| 8b. Drinking | 0-3 | 0.51 | (0.72) | 0.41 | 0.46 | | | (0.82) |
| 8c. Enteral Feeding | 0-4 | 0.65 | (1.46) | 0.56 | 0.61 | **0.56** | (0.56) | |
| 9. Skin Pressure | 0-5 | 0.73 | (1.39) | 0.67 | 0.72 | **0.73** | | |
| 10. Safety Awareness | 0-3 | 0.73 | (0.91) | 0.68 | 0.72 | | **0.76** | |
| 15.Communication | 0-5 | 1.03 | (1.48) | 0.47 | 0.52 | | **0.78** | |
| 16. Behaviour | 0-5 | 0.55 | (1.01) | 0.34 | 0.39 | | **0.64** | |
| SPECIAL NURSING NEEDS | | % positive | | | | | | |
| Tracheostomy | 0-5 | 4 | | 0.24 | | | | |
| Wound dressings | 0-5 | 10 | | 0.21 | | | | |
| >2 night time interventions | 0-5 | 22 | | 0.40 | | | | |
| Psychological support | 0-5 | 19 | | 0.18 | | | | |
| Isolation for MSRA | 0-5 | 15 | | 0.31 | | | | |
| Intercurrent medical/surgical condition | 0-5 | 7 | | 0.24 | | | | |
| 1:1 specialing | 0-5 | 2 | | 0.01 | | | | |

*All item ranges represented the full possible range of the score*

***Note: Item-factor loadings rounded to two decimal places and for the Varimax rotation, all loadings <.50 were removed for clarity*

## Factor analysis

*Table 20* also presents the results of the principal component analysis of the 16 BCN items. These indicate the presence of a large general 'nursing dependency' factor with 13 out of the 16 items loading high (i.e. >0.50) on the first principal component. Only three items had a loading below 0.5 (*eating*, *drinking* and *behaviour*). Three components had eigenvalues >1.0 and the parallel analysis also suggested a three-factor solution, accounting overall for 70% of the total variance. However, the third factor appeared to only reflect the local dependency of the eating and drinking items. (Local dependency refers to items with very high correlations between them due to the fact that a person cannot score high on one without also scoring high on the other). In this case, people who need help drinking virtually always need help eating, so this left a two-factor solution as shown in *Table 20*.

The first factor represents *physical care needs* (13 items) and accounted for 50% of the variance, the second factor represents *cognitive/behavioural needs* (3 items) and accounted for 12% of the variance. Coefficient alphas for these factors were 0.93 for the *physical care needs* subscale and 0.68 for the *cognitive/behavioural needs* subscale.

## Test Discrimination and Responsiveness

The ability of the BCN section to discriminate among people with different degrees of nursing dependency, as measured by Coefficient δ, was high at 0.99. *Table 21* shows the effect sizes for the NPDS scores in comparison with the BI, FIM and FIM+FAM for the four different dependency groups. As expected, responsiveness was greatest for the higher dependency patient groups, at which the NPDS is targeted.

**Table 21: Effect sizes (ES = mean change from baseline/SD baseline) and Wilcoxon Signed Ranks Z for the various scales within the different dependency groups**

| Dependency Scale | Low (NPDS 0-9) N=147 | | Medium (NPDS 10-24) N=183 | | High (NPDS 25-40) N=84 | | Very high (NPDS 41+) N=72 | |
|---|---|---|---|---|---|---|---|---|
| | ES | Z | ES | Z | ES | Z | ES | Z |
| **NPDS** | | | | | | | | |
| BCN | 0.58 | -6.55 | 1.79 | -10.70 | 2.19 | -7.49 | 1.89 | -6.51 |
| SNN | 0.37 | -2.53 | 0.25 | -3.03 | 0.57 | -3.99 | 1.03 | -6.00 |
| **Total** | 0.37 | -4.22 | 1.53 | -10.59 | 3.30 | -7.26 | 1.98 | -5.59 |
| **RCH** | 1.00 | -9.12 | 1.55 | - 10.83 | 1.52 | -6.34 | 2.41 | -5.69 |
| **Barthel** | 1.12 | -9.68 | 2.11 | -11.54 | 2.01 | -7.45 | 2.69 | -6.06 |
| **FIM** | | | | | | | | |
| Motor | 1.03 | -10.25 | 1.91 | -11.63 | 1.99 | -7.71 | 2.74 | -6.30 |
| Cognitive | 0.42 | -7.28 | 0.52 | -8.96 | 0.45 | -5.89 | 0.59 | -6.01 |
| **Total** | 1.04 | -10.31 | 1.87 | -11.66 | 1.76 | -7.71 | 1.85 | -6.83 |
| **FIM+FAM** | | | | | | | | |
| Motor | 1.20 | -10.43 | 2.09 | -11.63 | 2.11 | -7.74 | 2.41 | -6.44 |
| Cognitive | 0.12 | -1.43* | 0.40 | -6.80 | 0.36 | -4.88 | 0.63 | -6.35 |
| **Total** | **0.86** | **-9.59** | **1.60** | **-11.63** | **1.31** | **-7.54** | **1.31** | **-6.90** |

*[1.] Effect sizes can be interpreted according to Cohen (0.2=small, 0.5=moderate, 0.8=large).*
** = non-significant, all other Z values p<.01.*
*BCN = Basic Care Needs subscale of the NPDS, SNN = Special Nursing Needs subscale of the NPDS,*
*RCH = estimated nursing care hours, FIM = Functional Independence Measure.*

As shown in Table 22, The concurrent correlations of the different components of the NPDS with admission scores from the FIM/FAM and the Barthel provide further support for the validity of the measure. For example:

- The BCN score correlates with the Barthel, FIM Motor and FIM+FAM Motor at r=-0.87 to -0.88.
- The SNN score also shows significant correlations with each of these three measures but the correlations are moderate (r=-0.50 to -0.51) reflecting the more specialised nature of the SNN items.
- As expected, the NPDS cognitive items correlated more strongly with the FIM and FIM+FAM cognitive scales (r=0.74-0.77) than with the motor subscales or BI (r=0.49-0.53).

Parametric and non parametric tests produced very similar results.

### Convergent and discriminant validity

**Table 22: Pearson correlations of NPDS total score, BCN and SNN sections and NPDS cognitive item admission scores, with admission and discharge scores on BI, FIM, and FIM+FAM (Spearman rho in parentheses)***

| Measure | NPDS Admission | BCN Admission | SNN Admission | NPDS Cog Admission | RCH Admission | RCH Discharge |
|---|---|---|---|---|---|---|
| BCN admission | 0.97 (.97) | | | | | |
| SNN admission | 0.78 (.71) | 0.59 (.56) | | | | |
| Total NPDS Discharge | 0.81 (.77) | 0.80 (.77) | 0.60 (.51) | | | |
| **Barthel Index** | | | | | | |
| Admission | -0.83 (-.86) | -0.87 (-.90) | -0.50 (-.50) | -0.49 (-.51) | -0.85 (-.85) | -0.73 (-.72) |
| Discharge | -0.79 (-.77) | -0.81 (-.78) | -0.53 (-.44) | -0.53 (-.49) | -0.68 (-.70) | -0.88 (-.84) |
| **FIM Motor** | | | | | | |
| Admission | -0.84 (-.87) | -0.88 (-.90) | -0.50 (-.50) | -0.50 (-.51) | -0.86 (-.85) | -0.74 (-.73) |
| Discharge | -0.81 (-.81) | -0.82 (-.78) | -0.53 (-.44) | -0.54 (-.50) | -0.68 (-.71) | -0.88 (-.85) |
| **FIM Cognitive** | | | | | | |
| Admission | -0.60 (-.52) | -0.63 (-.55) | -0.35 (-.30) | -0.77 (-.74) | -0.45 (-.42) | -0.52 (-.47) |
| Discharge | -0.62 (-.52) | -0.63 (-.60) | -0.40 (-.32) | -0.74 (-.70) | -0.45 (-.41) | -0.60 (-.54) |
| **FIM+FAM Motor** | | | | | | |
| Admission | -0.85 (-.87) | -0.88 (-.90) | -0.51 (-.50) | -0.50 (-.52) | -0.85 (-.85) | -0.74 (-.73) |
| Discharge | -0.82 (-.80) | -0.83 (-.82) | -0.54 (-.45) | -0.55 (-.50) | -0.70 (-.73) | -0.89 (-.86) |
| **FIM+FAM Cognitive** | | | | | | |
| Admission | -0.67 (-.60) | -0.70 (-.63) | -0.40 (-.34) | -0.77 (-.76) | -0.53 (-.51) | -0.58 (-.54) |
| Discharge | -0.66 (-.58) | -0.68 (-.60) | -0.43 (-.32) | -0.76 (-.73) | -0.49 (-.57) | -0.65 (-.61) |

** Note: N ranges from 490–565. All correlations are significant p<0.01 (2-tailed)*

Together with the systematic review, the findings from this analysis demonstrate that there is now a substantial body of evidence for the psychometric properties of the NPDS, which is summarised in *Table 23,* against the Medical Outcomes Trust criteria.

**Table 23: Psychometric evaluation of the Northwick Park Dependency Scale according to the Medical Outcomes Trust framework**

| Attribute and criteria | Evaluation |
|---|---|
| ***Conceptual and measurement model:** The rationale for and description of the concept and the populations that the measure is intended to assess* | |
| **Clinical content and design** | A 23-item ordinal rating scale to quantify an individual's needs for nursing care and support – particularly highly dependent patients.<br>Subscales: Basic Care Needs (BCN) 16 items- range 0-65) and<br>Special Nursing Needs (SNN) 7 items (0-35). Total score range 0-100<br>Translates by a computerised algorithm into an assessment of care hours, and costs of care in the community – the NPCNA.[77] |
| **Dimensionality** | Principal component analysis demonstrates a strong general 'dependency' factor within the BCN subscale, with two major specific factors ('Physical Care needs' and 'Cognitive/Behavioural needs'). |
| ***Reliability:** The degree to which the instrument is free from random error* | |
| **Internal consistency and homogeneity** | Cronbach's alpha: For the full 23-item NPDS scale (alpha = 0.90); for the 16-item Basic Care Needs scale (alpha = 0.93). For the 7-item Special Nursing Needs scale alpha = 0.50 (reflects lower positive score rates for these items).<br>Item-total correlations: BCN section: 0.34 – 0.88; SNN section 0.18 – 0.31 (also reflects lower positive score rates for these items). |
| **Reproducibility** | Inter-rater reliability: Three studies report reliability correlations (>0.80) for the full NPDS score and the BCN scale.[77, 88, 89]<br>2 studies reported correlations for the SNN scale (0.48 - 0.80).[77, 88]<br>Agreement: Three studies reported absolute agreement for individual items (39-100%) ; unweighted kappa 0.63 – 1.0.[77, 88, 89] |
| ***Validity:** The degree to which the instrument measures what it purports to measure* | |
| **Content** | Within developmental design – based on clinicians' expert opinion, and several rounds of observed activity analysis[90] and work sampling in a hospital setting.[91] |
| **Criterion-related** | Not testable - no accepted gold standard currently exists. |
| **Construct** | Five studies reporting high correlations (0.83-0.95) between BCN and/or total NPDS scores with the BI[77, 88, 92, 93]; and between the BCN and FIM.[89] Moderate correlation between SNN scale and Barthel (0.50). |
| ***Discrimination and responsiveness:** Ability to detect change over time where real changes occur* | |
| **Change: admission to discharge** | Discriminates among people with different levels of dependency (Coefficient δ 0.99). Responsive to change over time, particularly in the higher dependency groups (effect – size 1.9 – 3.3). |
| ***Interpretability:** Whether easily understood meaning can be assigned to quantitative scores* | |
| **Clinical meaning** | Nurses preferred the NPDS over the BI because it provided 'better, information about the actual need for care of the patients'.[92] |
| ***Burden:** The time, effort or other demands of administering the instrument* | |
| **Time to administer** | The mean time to complete the NPDS for an experienced user was under five minutes (n=22 patients).[88] |
| **Alternative modes of administration** | A self-report version for patients and their carers is currently undergoing testing. |
| **Cultural and language adaptations** | Swedish[89] and Dutch.[93] |

** RCS T score = Therapy Disciplines (TD) +Therapy Intensity (TI)*

Our findings were thus generally supportive of the reliability and validity of the NPDS as a measure of nursing dependency for use in rehabilitation settings. High internal consistency and the presence of a large principal component within the BCN scale, support its use as a uni-dimensional scale for measuring overall need for nursing support. However, the factor analysis demonstrated that the BCN scale can usefully be considered as having two substantive subscales, one reflecting needs for physical care, and the other cognitive/behavioural needs. The NPDS also showed excellent discriminatory power for people with different degrees of dependency, and was responsive to change - particularly for the medium and high dependency groups of patients, where the effect sizes ranged from 1.5 – 3 standard deviations.

Comparison with the BI and FIM confirmed a close relationship between the three scales based primarily on the similarity of these other two scales with the BCN section of the NPDS. The physical and cognitive items of the BCN mapped broadly onto the motor and cognitive elements of the FIM+FAM, confirming that the NPDS provides information on cognitive or behavioural problems in addition to detailing the nursing support that is needed for physical functioning – and in this respect it has a modest advantage over the Barthel Index. High correlations between the admission NPDS and discharge FIM and BI suggest that the NPDS admission scores may also have predictive value as an indicator of outcome, in similar fashion to the FIM.[94]

One of the key differences between the NPDS and these other scales is the additional SNN section which describes the needs for more specialised nursing support such as tracheostomy, wound care or one-to-one nursing. Evidence from observational studies confirms that, when these needs arise, they have considerable impact on nursing time and skills,[90] and these are important for the purposes of planning staffing provision in relation to caseload, as they are generally tasks that require input from a qualified nurse.[85] We expected, and found, a weaker relationship between item and total scores for this section of the scale, and also with the BI and FIM confirming that it is indeed measuring something different. Even in this large and relatively complex sample, the item frequencies were low for some items in this section (notably 1:1 special nursing), which will have affected the overall item total-correlations.

A further point of potential added value for the NPDS is its direct translation into an assessment of care hours. The BI and FIM are both shown to correlate with care hours on a population level,[85, 90] but cannot be used to measure them directly as they do not assess the number of people required to assist with the task nor the time taken to complete it. Part of the potential added value of the NPDS is its algorithm to calculate the impact of reduced dependency on care hours and costs at the level of the individual. This may be one of the reasons for the growing popularity of the NPDS within the UK, but further research is still required to define the differential calculations for care hours in different settings and to evaluate them internationally in different health cultures.[85, 90]

The study also highlighted that the SNN section is both conceptually and structurally different from the BCN. Given the strong correlation between the two subscales (0.56, $p<0.01$), it appears reasonable to sum the two sections to yield a total NPDS score as an overall indicator of how dependent the person is on nursing support. However, on a practical clinical level it is appropriate to present the scale totals separately.

In summary, there was strong evidence in support of the NPDS as a valid and reliable measure of dependency for basic care and nursing, but more work was required to render it suitable for the different settings in which it would be used in the course of the programme.

In addition, we note parametric and non-parametric analytical techniques gave very similar results in this evaluation. This provides supportive evidence for parametric analysis being an acceptable approach for long ordinal data of this kind in reasonably sized sample populations, even if the data are not normally distributed.

## *Translation into care hours and costs*

As noted above, a critical feature of the NPDS is that it translates by a computerised algorithm into the Northwick Park Care Needs Assessment (NPCNA). This provides estimations of the care hours per week, and the approximate cost per week of the care package that would be required to meet the patient's needs for basic care and nursing in the community. The NPDS has been used to evaluate the cost-efficiency of rehabilitation by comparing the cost of care/week in the community before and after rehabilitation, to determine how long it would take to offset the initial costs of the rehabilitation programme through on-going savings in the cost of community care.

A single centre study published in 2006/7[10] showed that specialist rehabilitation was cost-efficient, especially for highly dependent patients in whom the costs of rehabilitation were offset within just 16 months of discharge. It also demonstrated clearly that longer-stay rehabilitation programmes can provide value for money.[11]

This evidence has proved to be compelling in terms of policy development, and has helped to convince commissioners of health care services (NHS England) and Clinical Commissioning Groups (CCGs) of the need to invest in specialist rehabilitation services.[95] However, the argument rests on an assumption that dependency remains at least stable (or even continues to fall) over a sufficient number of months to recoup the financial benefits. At the outset of this programme, that evidence was missing.

Part of this programme was conducted in parallel with a project funded by the NIHR Services Delivery and Organisation (SDO) programme.[96] This study followed-up patients admitted to any of the nine Level 1 services in London for up to 1 year after discharge to evaluate the extent to which their needs for ongoing-rehabilitation and support were provided for, and to relate met and unmet needs to longer term outcomes.[97] The NPDS /NPCNA formed part of the dataset for that follow-on study, and this provided an opportunity to examine how far the gains in reduced dependency and care costs were sustained in the 12 months after discharge from in-patient rehabilitation.

As the recruitment period for this study pre-dated the UKROC database, the NPDS/NPCNA did not form part of mandated data collection at the time of recruitment, so the community follow-up data could only be linked to inpatient data from one centre (the Regional Rehabilitation Unit, Northwick Park Hospital) where the NPDS was already consistently recorded for all inpatient admissions. This provided a cohort of 81 consecutive patients – the demographics of this cohort is shown in *Table 24.*

**Table 24: Demographic profile of patients participating in the community follow-up study**

| | Total recruited sample 2009 study (N=81) | Sub-sample responding at all time-points (n=39) |
|---|---|---|
| **Males: Females n (%)** | 64:36% | 62:38% |
| **Age (Mean SD)** | 44.3 (13.7) | 45 (13.1); |
| **Length of stay (Mean SD)** | 106.5 (62) | 95 (54 days) |
| **Cost of rehabilitation (Mean SD)** | £51,342 (£29,863) | £45,864 (£26,992). |
| **Diagnostic group n (%)** | | |
| Acquired brain injury | 68 (84%) | 33 (85%) |
| Spinal cord Injury | 5 (6%) | 2 (5%) |
| Other | 8 (10%) | 4 (10%) |
| Missing | 0 | 0 |
| **Admission (Mean SD)** | | |
| **NPDS total dependency score** | 21.5 (15) | 22.1 (17.3) |
| **Restricted weekly care hours** | 40.1 (20) | 38.2 (19.4) |
| **Average weekly cost of care** | £1,184 (£1,025) | £1,085 (£918) |

The mean NPDS score was 21.5 (SD 15) on admission, which translated into 40.1 (SD 20) care hours per week, at an average cost of £1,184 (SD £1,025) per week.

*Table 25* shows the change in these parameters from admission to discharge. With a mean cost of the inpatient rehabilitation programme at £51,342 and a mean saving of £470 per week, the costs of rehabilitation in this group was offset within 27 months.

**Table 25: Changes in dependency, care hours and costs from admission to discharge (N=81)**

| Change over time admission to discharge (N=81) | Mean difference | SD | 95% CI | t-value | p-value |
|---|---|---|---|---|---|
| **NPDS total dependency score** | 9.6 | (9.3) | 7.6, 11.7 | 9.3 | <0.0001 |
| **Restricted weekly care hours** | 13.0 | (13.4) | 10, 15.9 | 8.7 | <0.0001 |
| **Average weekly cost of care** | £470 | (£563) | £345, 594 | 7.5 | <0.0001 |

Follow-up evaluations were achieved through questionnaires supported by telephone interviews at 1, 6 and 12 months post discharge. Follow-up for this group of patients is known to be challenging,[98] and there was notable attrition of the study sample over the 12 month period. The number of patients responding at the different time points varied somewhat but n=43 provided data at 12 months. The distribution of NPDS scores is shown in *Figure 19*. There was a slight increase in dependency from discharge to month 1, which was not unexpected as patients adjusted to their new environment, but by 12 months dependency scores had dropped back to the same level as at discharge.

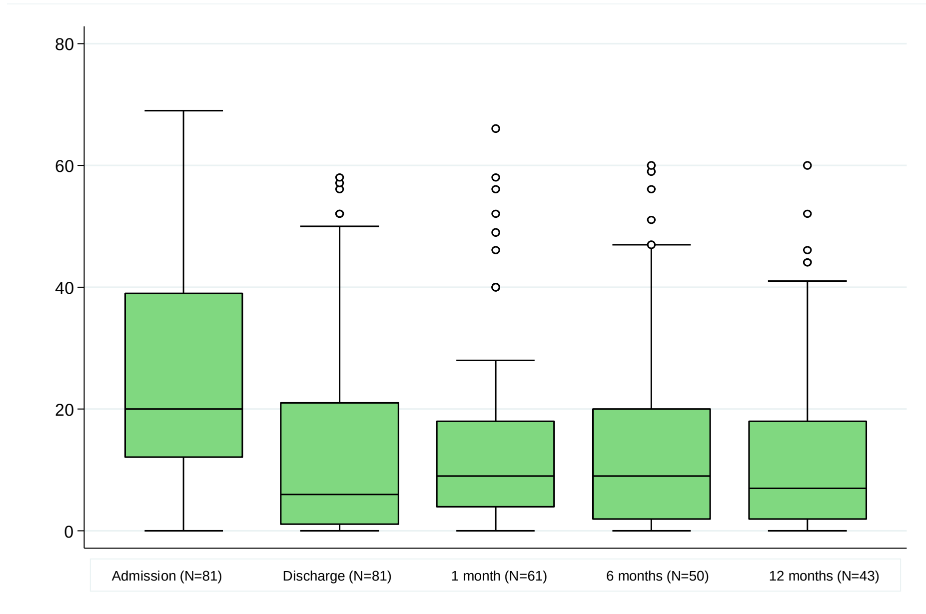



**Figure 19: Serial box plots showing change in dependency over time**

As the precise group of patients responding at the different time points varied somewhat, we could not exclude a sampling error. Therefore we undertook an analysis of the 39 patients (48% of the sample) who responded at every time-point. The demographics of this group were not significantly different from the total sample population (N=81) as shown in *Table 24.*

*Figure 20* shows the NPDS scores and NPCNA-estimated costs of care per week for this sample across all five-time points. Once again the same pattern is seen with a marked reduction in both parameters from admission to discharge, followed by slight increase at the 1month post-discharge point but settling back to the discharge levels or below by 12 months.

As the numbers were small and the data not normally distributed, we used non-parametric statistical techniques.  Wilcoxon signed rank tests confirmed significant differences only between the admission and discharge data for NPDS (z=4.6, P<0.001) and care costs/week (z=4.3, p<0.001).  No significant differences were seen for any other time-points after discharge.

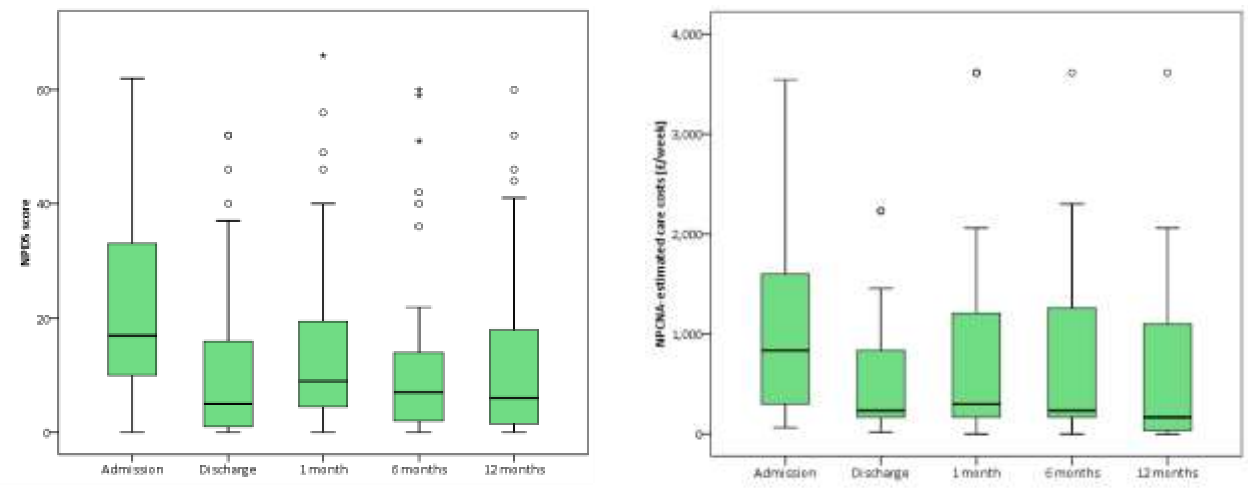



**Figure 20: NPDS scores and NPCNA-estimated costs of care per week for the complete responders (N=39)**

In summary, these data provide some evidence that gains achieved in the reduction of dependency and care costs during inpatient admission are likely to be sustained over periods of at least a year. It supports the use of time to offset the initial costs of rehabilitation as a valid measure of cost-efficiency.

Weaknesses of the study were:

a) That less the half of the patients responded at every time point (although this group appeared to be representative of the larger group, as they did not differ significantly in any sense).

b) Follow-up was only to 12 months, rather than the full 27 months required to offset the cost of rehabilitation. On the other hand, the trend if anything was towards further improvement over the 12 month period.

## *Further development of the NPDS for hospital settings*

The original conceptualisation and development of the NPDS was focussed on deriving care needs in the community.  Although an individual's care needs may be very similar whether they are in hospital or in the community, there are several differences between hospital and community settings which may affect precise estimation of staffing hours required to support those needs.  For example, in the community, nursing and care staff generally care for one patient at a time and need to travel between them, whereas in a ward setting, a team of nurses is present at all times, and a single nurse can often supervise two or more patients simultaneously if they only need prompting or incidental help for certain elements of care.  If the NPDS were to be used to determine ward staffing levels, it was therefore necessary to consider how the NPDS/NPCNA could be used directly, or alternatively adapted for this purpose.

In work published just prior to the start of this programme in 2007, we compared NPCNA-derived estimates of the total care-hour requirements for the caseload at any given time, with the total nursing staff hours available from the duty roster.[99]

- Over a 6-month period, there was a modest relationship between these two parameters (Pearson's r=0.31); but staff hours provided fell significantly short of estimated daily care requirements (mean shortfall 6.2 hours/day, SD 8.6 (95% CI 0.73, 11.1, p <0.03).
- The discrepancy between staffing levels and care needs varied throughout the day. The shortfall was most obvious during peak care periods in the morning and at bedtime, which resonated with experience reported by nursing staff that they are overstretched at these times.
- A marked discrepancy between qualified nursing staff hours provided and NPCNA-estimated 'special nursing' care requirements highlighted a failure to address important facets of the rehabilitation nursing role.

We concluded that, whilst the NPDS could continue to be used to assess dependency, a different algorithm was required to calculate nursing staff hours in order to determine staffing levels required to meet patients' care needs in the inpatient setting.

Work to underpin development of an algorithm to calculate staff hours from the NPDS in inpatient settings included:

A. In a qualitative study, a series of thee focus groups with experienced rehabilitation nurses/ health care assistants to examine the range of activities undertaken by nursing staff in an rehabilitation setting compared with community settings.

B. A timed observational study[100] to record the time taken to complete a range of individual nursing care interventions in a rehabilitation ward.

C. A work sampling study to identify the proportions of staff time dedicated to direct patient care and other (indirect or non-patient-related) activities in an in-patient rehabilitation setting.[101]

A and B pre-dated this programme and C was completed within it.

The qualitative study (A) concentrated on changes to the NPDS to make it more suitable for in-patient settings. The principal change suggested by the focus groups was to convert the SNN items from a simple dichotomous rating scale to an ordinal scale, with the aim of providing greater sensitivity to small changes. In the timed observational study (B), a total of 1168 nursing interactions were timed for 50 care episodes. The principal conclusions from this study were as follows:

- The total nursing time taken to provide direct care for an individual patient varied with dependency, as we expected. Median (IQR) times for care during the 18-hour observation period (6am to midnight) were:
    - 22 minutes (IQR 11–52) for the low dependency group (NPDS<10),
    - 106·5 minutes (IQR 68·5–135·5) for the median dependency group (NPDS=10-25),
    - 293 minutes (IQR 258–389 minutes) for the high dependency group (NPDS>25).
- The differences were highly significant between all groups *(*p< 0·001).

The NPCNA-estimated care times were significantly greater than the observed care times for interventions, so the latter were used to develop a hospital-based algorithm. In addition, in the hospital version of the NPDS (the NPDS-H), the SNN section was expanded to accommodate the additional range of interventions observed in hospital, and instead of the dichotomous scoring (0 or 5) an ordinal scale was introduced for each item. For the hospital version of the tool (the NPDS-H), see *Appendix 4.8.*

By the start of this programme, therefore, we had a set of tools for estimating the care/nursing hours required for direct hands-on patient care for use both in community and hospital rehabilitation settings.

However, nurses also do many other things during a nursing shift.  In order to apply this information to estimate the actual staffing requirements in relation to a given rehabilitation ward caseload, it is necessary to know the proportion of time nurses spend in other activities and how this may vary throughout the day. It is also pertinent to know which tasks require a registered nurse and which can be done by non-registered staff such as health care assistants (HCAs) and student nurses.

In the early stages of this programme we analysed data from C, the work sampling study, which had been gathered in 2004, but not previously published.  The full methods and results may be found in Williams et al 2009,[101] but are summarised briefly here.

The study involved independent observation of activities undertaken by all nursing staff on the unit at five-minute intervals over a period of two weeks.  It incorporated a total of 126 hours during which 8883 five-minute episodes of nursing activity were observed and recorded.  During this period there were 23 patients cared for by 34 members of staff (13 registered nurses and 21 non-qualified health care assistances). Activities were broken down into direct and indirect patient-related care, other unit related activities, personal time (rest-breaks, appraisal continuing profession development etc).

Whilst patient-related care (direct and indirect) accounted for 71% of the nursing activities overall, the proportion of time spent on these varied through the day (being greatest in the morning and evening as patients are got up and dressed, and then settled back in bed for the night).  This pattern was also reflected in the NPCNA estimated hours, see *Figure 21*.

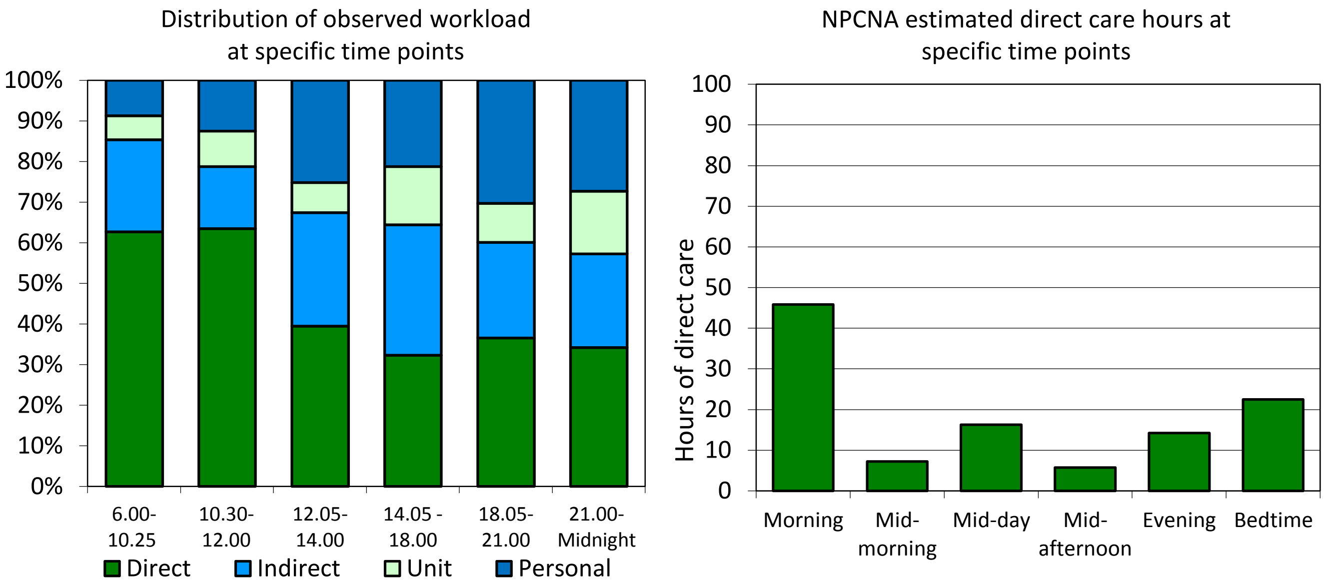



**Figure 21: Comparison of trend in direct care allocation between observed activity and estimated community care hours using Northwick Park Care Needs Assessment (NPCNA)**

The distribution of observed nursing workload is shown in *Table 26*.

- Direct and indirect patient care accounted for 75% of RN time and 68% of HCA time.
- RNs were more involved in indirect patient care (e.g. goal setting, team meetings telephone liaison and patient documentation), while the HCAs spent more time on direct patient care.
- Both spent a spent a similar proportion of their time on with ward-related duties, but HCAs spent a greater proportion of their time on personal activities (including training).
- The overall NPCNA estimated care hours were approximately twice the observed hours spent in direct patient care, but they correlated very strongly (rho 0.86 p<0.001).

**Table 26: Distribution of observed nursing workload**

| | Registered Nurse, N = 3386 (% of RN activities) | Non-registered staff, N = 5497 (% of non-RN activities) |
|---|---|---|
| **Direct patient care** | 1253 (37%) | 2807 (51%) |
| **Indirect patient care** | 1273 (37%) | 945 (17%) |
| **Unit-related** | 327 (10%) | 547 (10%) |
| **Personal time** | 533 (16%) | 1198 (22%) |

*RN, Registered Nurse.*

- However direct care accounted for just under half of the total staff time when indirect care and nonclinical activities were taken into account.
- This suggests that, once the other activities are taken into account, the NPCNA-estimated care hours (RCH) may give a crude overall estimation of the total time requirements for nursing/care staff, which will vary with the complexity of the caseload.
- Whilst it may not be accurate for absolute estimation of staff time requirements, it was felt to be sufficiently representative to estimate the relative proportion of staff time spent on patients with different levels of dependency, pending the development of a more accurate algorithm.

The NPDS-H was introduced and incorporated within the UKROC software from 2011. However, many units were more familiar with, and therefore still recording the original NPDS. In order to maximise the generalisability of the data, we therefore opted to use the original NPDS and NPCNA algorithm for our calculations of the relative proportions of staff time when deriving the proportionate costs of different levels of complexity (see *Chapter 7*).

# The Northwick Park Therapy Dependency Assessment (NPTDA)

## *Introduction*

A substantial literature now supports the benefits of higher intensity rehabilitation, at least for certain patients,[102-104] but 'higher intensity' has yet to be properly defined. Patients with neurological disabilities have widely varying needs for rehabilitation, often involving several disciplines. Simply recording hours of therapy input has little meaning unless the nature of interventions can be also be described. Many authors have called for practice-based research to 'open the black box', in order to provide clearer description of the rehabilitation content.[105] A number of tools have been developed to facilitate the systematic recording of therapy interventions,[106-113] which include tools to describe the type of interventions offered for patients with stroke[106, 108-111] and spinal cord injury.[112, 113] However, these can only be applied to describe interventions that were *actually given*, rather than looking at what might be *needed*. Moreover, existing tools focus only on physical interventions (physiotherapy, occupational therapy and in some cases speech and language therapy[109]) and omit other interventions such as psychology, dietetics and social work, which play an important role in holistic neurological rehabilitation programmes.

In 2004, a project grant was awarded by the UK Department of Health (Grant ref 030/0066) to develop an equivalent tool to assess therapy dependency.[53] The Northwick Park Therapy Dependency Assessment (NPTDA) was developed through an iterative process over two years. However, at the outset of this programme it had not yet been published.

In the early stages of this programme we published a paper describing its development and initial validation.[114] As this paper is now in the public domain, the findings are only summarised briefly here.

## *The NPTDA*

The NPTDA is a measure of therapy intervention designed for use in specialist neuro-rehabilitation settings, where high intensity rehabilitation is provided by a MD team. The tool is designed to be used either prospectively (to estimate anticipated therapy needs), or retrospectively (to quantify therapy actually provided). The difference between these two can then be used as a measure of *met* and *unmet* need for therapy interventions.

Key principles of the tool are as follows:

- It includes 30 items of therapy dependency in seven domains (A-G), see *Appendix 4.7a*. For the general scale structure, see *Appendix 4.7b*. The total range of the score is 0-100.
  - Items in domains A-E record direct 'hands-on' patient care, each scored on a range of 0-4.
  - Items in domain F record indirect patient-related care (e.g. attending meetings, writing reports etc which may be conducted away from the patients), and additional activities such as groups or staff-escorted clinic attendance. These items are scored on a range of 0-2.
  - Items in domain G are 'text only' and record the use of special facilities/equipment, investigations and procedures, for the purpose of audit and coding.
- Each patient is rated individually, the scores for each item being based on the interventions for a one-week period.
- A scoring manual provides detailed level descriptions for each item.
- The data are entered into an electronic database, which applies a computerised algorithm to estimate the therapy hours for each level of each item, see *Appendix 4.7c.*
  - For score levels 1-3, the algorithm applies pre-determined hours, which are allocated to the lead discipline identified. A default lead discipline is suggested for each item, but this may be changed to reflect normal practice within a given setting.
  - Levels 3.5 and 4 reflect interdisciplinary working, where several different disciplines are working in collaboration on the same task area (item). In this case, the hours are specified individually for each discipline on the scoring sheet at the time of rating.
- The allocated times are summed to provide an estimate of the total therapy hours and also provide a breakdown of hours for each discipline.

## *Initial Validation*

The initial validation was conducted in two stages:

In the first stage, we validated the NPTDA scores, and refined the conversion algorithm for translating raw scores into therapy hours, by comparing retrospectively-applied NPTDA estimates of therapy intervention with the actual hours of therapy intervention - recorded through parallel systematic activity analysis.
As well as recording the 'actual hours' per item for each patient, we also mapped these by reverse transcription to derive NPTDA scores from the activity analysis ('Activity analysis-derived NPTDA' scores), using the time range stated for each item level within the scoring manual. Agreement between 'actual' and 'derived' scores was tested using Cohen's Kappa.

- NPTDA-estimated therapy hours/week were strongly correlated with those identified from activity analysis, for total scores (Spearman rho 0.77, $P<0.0001$), and also for all five sub-domains for direct (hands-on) intervention (rho 0.70-0.93, $p<0.0001$).

- Nine of the 22 direct intervention items achieved 'substantial' or 'almost perfect' agreement (weighted kappa > 0.65), a further eight achieved moderate agreement. Four direct intervention items and two indirect intervention achieved only fair agreement (kappa 0.2-0.4).
- The initial test algorithm over-estimated therapy hours (Wilcoxon z-3.9 p<0.001). After adjustment, re-analysis using a revised algorithm showed this bias to be removed (Wilcoxon z=1.4 p=0.15).

In the second stage, using a subsequent cohort of patients, we compared prospective and retrospective NPTDA ratings, recorded in parallel for the same treatment period, to examine the validity of prospective application.

- Prospective and retrospectively-applied total NPTDA scores were strongly correlated (rho 0.61, p<0.0001).
- Although intended levels of intervention were higher than those actually delivered (Wilcoxon z=-3.30, p<0.001), the differences corresponded to real deviations from intended practice.

**Strengths and limitations:**

The study captured activity for approximately 420 therapist-days over the study period (20 working days for 20.3 WTE) generating a large quantity of data. We recognised the potential for rating bias and incomplete capture of activities, and despite the large volume of data it was a relatively small sample from a single unit with a high proportion of acquired brain injuries.

Nevertheless this study provided preliminary evidence for the content and face validity of the NPTDA as well as a validated algorithm for translation into therapy staff time. The next step was an evaluation of the construct and concurrent validity of the NPTDA in a wider more general neurorehabilitation sample.

## *Further evaluation of the NPTDA within this programme*

Different countries have widely differing health cultures. It is not expected that an algorithm developed in the UK for translating the NPTDA into staff time would necessarily be applicable internationally. However, the simple ordinal NPTDA potentially fills a gap in the international toolset for measuring therapy inputs. In countries such as the United States and Australia, case-mix systems use the Functional Independence measure (or FIM™ (Uniform Data Systems)) as a proxy for therapy needs.[12] However, for patients with complex neurological disabilities, physical dependency is not necessarily a good indicator of needs for therapy intervention.[13] Moreover, it is expected that therapy interventions will change over the course of a rehabilitation programme, not only in the quantity (or intensity) of input required, but also in the focus for intervention. For example, in the early stages following a severe stroke or brain injury, much of the focus for intervention may be on restoring physical function, managing basic needs such as nutrition and tracheostomy weaning and working towards independence in basic self-care. Towards the end of the programme the focus is expected to change towards discharge planning and community re-integration. It is therefore relevant to have a tool that is practical to apply serially over time, which captures both the quantity and nature of therapy interventions provided by the whole multi-disciplinary team.

Accepting that rehabilitation services outside of the UKROC system might wish to use the NPTDA as a simple ordinal scale, without necessarily applying the algorithm to generate therapy hours and convert it to an interval level measure, it was appropriate to explore the scaling properties of the basic ordinal tool.

We therefore examined its psychometric properties in terms of factor structure, internal reliability and responsiveness in a large mixed multicentre neurorehabilitation sample. We also assessed concurrent validity through exploring its relationship with the UK-ROC needs and outcome scales.

This paper has been published in Disability and Rehabilitation 2015,[115]but an abbreviated version is included here.

## Methods

### Data source

The UK Rehabilitation Outcomes Collaborative (UK-ROC) database was established in September 2009. Formal data collection started from January 2010, initially on a voluntary basis. To assess the psychometric properties of the NPTDA scale, we extracted the consecutive cohort of all patients admitted to specialist rehabilitation services within England reporting data to UKROC between 1.1.10 and 30.11.12, for whom an NPTDA score was available at admission and/or discharge from hospital. The dataset comprised 2505 patients in 49 neurorehabilitation centres (representing approximately 82% of the total number of services reporting data during the study period).

### Measures

In addition to the NPTDA, other measures used in this analysis were the RCS (version 8) and the FIM component of the UK FIM+FAM. These were used to describe the population in terms of overall complexity of rehabilitation needs and levels of functional independence at admission and discharge for rehabilitation. There is no gold standard against which to determine the validity of the NPTDA, but its relationship with complexity and functional independence was used to provide an overall indicator of concurrent validity. We used the FIM score because this is widely used and understood by rehabilitation professionals and is used in some health system as a proxy for rehabilitation needs. We did not expect to find a close relationship because, as noted above, physical dependency is not the only (or even the main) determinator of rehabilitation needs. Nevertheless we expected to find weak-moderate negative correlations between the FIM and the NPTDA scores.

### Data extraction and Analysis

For the purpose of factor analysis we required two similar samples that spanned the range of NPTDA scores in all items. As we anticipated that admission scores would be systematically different from discharge scores, the admission and discharge records were combined into one dataset, which was then randomly split into two halves for our exploratory and confirmatory analyses. The combined total set consisted of N=3921 scores.

After cleaning to delete all records with missing information on individual NPTDA items, 3764 scores remained in the sample (2017 admission and 1747 discharge scores), equating to a 4.1% loss of records. This dataset was randomly split into two halves, each of n=1882 NPTDA scores. The first was used for the exploratory factor analysis and the second sample was used for confirmatory factor analysis.

Exploratory and confirmatory factor analysis were conducted using the methodology described earlier in this chapter. We used the 22 NPTDA ordinal scale items in domains A-D for the factor analysis. We excluded domains F and G, which are not designed to be scalable in the same way as the direct hands-on therapy items.

As the FIM, NPDS and NPTDA all generate ordinal data, non-parametric techniques should technically be used for comparing differences. On the other hand, factor analysis uses parametric assumptions. Given the large size of the dataset and distribution within acceptable limits of normality, parametric techniques (paired T tests) were used to describe differences between admission and discharge and p values <0.01 were considered statistically significant. *Table 27* shows the characteristics of the study population (n=2505 patients) and *Table 28* shows their scores on outcome measures at admission and discharge.

**Table 27: Characteristics of the study population (N=2505)**

| Characteristics | Mean (SD) | Range |
|---|---|---|
| Age, years, | 50.9 (15.9) | 14-91 |
| Length of stay | 83 (77) | 4-945 |
| | **N** | **%** |
| Males: Females | 1517:988 | 61:49 |
| **Diagnosis** | | |
| **Acquired brain injury** | **1602** | **64** |
| Stroke | 786 | 49.1 |
| Traumatic | 435 | 27.2 |
| Tumour | 122 | 7.6 |
| Anoxia | 101 | 6.3 |
| Inflammatory | 60 | 3.7 |
| Other | 98 | 6.1 |
| **Progressive Condition** | **242** | **9.7** |
| Spinal Cord Injury | 191 | 7.6 |
| Peripheral Neurology | 116 | 4.6 |
| Other | 173 | 6.9 |
| Unknown | 184 | 7.2 |

The population was slightly older than the group used in the initial evaluation, with a broader spread of neurological conditions. There were statistically significant differences between admission and discharge for the RCS and FIM scores (p<0.0001) in the expected directions. Complexity scores reduced and independence scores increased over the course of admission. However, no significant differences were observed between admission and discharge for the NPTDA overall total scores.

**Table 28: Scores on outcome measures at admission and discharge (N=2505)**

| | Admission | | Discharge | | Difference | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Rehabilitation scores | Mean (SD) | Range | Mean (SD) | Range | Mean | 95% CI | t | df | P value |
| **Therapy Dependency (NPTDA)** | | | | | | | | | |
| NPTDA total score | 18.1 (9.1) | 0-57 | 18.5 (9.9) | 0-55 | -0.4 | -0.8, 0.4 | -1.77 | 1415 | 0.077 |
| NPTDA-estimated total hours/week | 18.9 (16.4) | 0-227 | 18.4 (16.2) | 0-214 | 0.5 | -0.2, 1.2 | 1.33 | 1415 | 0.184 |
| **Rehabilitation Complexity Score (RCS-v8)** | | | | | | | | | |
| RCS Total | 10.4 (2.0) | 2-15 | 7.8 (2.9) | 0-15 | 2.6 | 2.5, 2.7 | 41.37 | 1885 | **<0.0001** |
| Care + Nursing + Medical | 5.3 (1.6) | 0-9 | 3.6 (2.1) | 0-9 | 1.7 | 1.6, 1.8 | 38.18 | 1885 | **<0.0001** |
| Therapy | 5.1 (0.9) | 0-6 | 4.2 (1.4) | 0-6 | 0.9 | 0.8, 1.0 | 28.01 | 1885 | **<0.0001** |
| **Functional Independence Measure (FIM)** | | | | | | | | | |
| FIM Total | 67.1 (30.6) | 18-126 | 89.5 (32.2) | 18-126 | -22.4 | -23.4, -21.3 | -42.46 | 1656 | **<0.0001** |
| FIM-Motor | 44.7 (24.9) | 13-91 | 63.1 (26.1) | 13-91 | -18.4 | -19.2, -17.4 | -40.58 | 1650 | **<0.0001** |
| FIM-Cognitive | 22.5 (9.8) | 5-35 | 26.6 (8.7) | 5-35 | -4.1 | -4.3, -3.8 | -28.19 | 1658 | **<0.0001** |

When the data were pooled and randomly split into two samples of n=1882, the first sample contained data points from a total of 1113 patients, and the second from 1148 patients. The groups were similar in terms of NPTDA total score (mean(SD), 18.7(9.5) versus 18.4(9.7), p=0.879), age, gender, length of stay and diagnosis.

*Table 29* shows the descriptive statistics for each item in the NPTDA scale based on the first randomly split sample (N=1882). We did not expect to find a homogeneous scale, but were interested to explore to what extent the domain structure developed intuitively during the scale design was borne out by factor analysis when the scale is applied in clinical practice. As expected, homogeneity was low for the full 26 item NPTDA scale with item-total correlations ranging from 0.15 to 0.49, and less than 50% being above 0.30. Nevertheless, Cronbach's coefficient-$\alpha$ was within the limits for good internal consistency at 0.76.

Principal component analysis of the 22 NPTDA items revealed 7 factors with eigenvalues >1.0. However, inspection of the scree plot showed a break after the fourth factor suggesting that a four-factor model was the best according to Gorsuch's criteria – these accounted for 43% of the total variance. The factors have been labelled as follows:

1. *Physical* (8items with 5 loading over 0.5),
2. *Psychosocial* (5 items, all with loadings over 0.5),
3. *Discharge planning* (5 items with 3 loading over 0.5) and
4. *Activities* (4 items with 2 loading over 0.5).

Although the item 'speech/language' loaded highest onto the Activities factor, it also loaded significantly (>0.3) onto the Physical factor, and it was assigned to the latter on the basis of best clinical fit.

**Table 29:** **Exploratory factor analysis: pattern matrix loadings for the hypothesised model (direct patient care)**

| Item (domain) | | Range | Mean (SD) | Item-Total Correlations | 1st Principal Component | Factor 1 Physical (17% variance) | Factor 2 Psychosocial (12% variance) | Factor 3 discharge planning (8% variance) | Factor 4 Activities (6% variance) |
|---|---|---|---|---|---|---|---|---|---|
| 1. | Medical (A) | 0-4 | 1.44 (1.09) | 0.26 | 0.43 | 0.54 | | | |
| 2. | Splinting (A) | 0-8 | 0.70 (1.32) | 0.33 | 0.43 | 0.45 | | | |
| 3. | Seating (A) | 0-4 | 0.99 (1.18) | 0.24 | 0.40 | 0.71 | | | |
| 4. | Physical therapy (A) | 0-4 | 2.58 (1.27) | 0.22 | 0.35 | 0.59 | | | |
| 5. | Tracheostomy (B) | 0-4 | 0.14 (0.60) | 0.13 | -0.39 | 0.49 | | | |
| 6. | Swallowing (B) | 0-4 | 0.33 (0.79) | 0.26 | 0.45 | 0.57 | | | |
| 7. | Nutrition (B) | 0-4 | 0.51 (0.94) | 0.28 | 0.46 | 0.66 | | | |
| 8. | Communication (B) | 0-4 | 0.39 (0.93) | 0.33 | 0.52 | 0.41 | | | |
| 9. | Speech/language (B) | 0-4 | 0.74 (1.11) | 0.38 | 0.54 | 0.35 | | | 0.50 |
| 10. | Personal/ self-care (C) | 0-4 | 1.11 (1.21) | 0.32 | 0.40 | | | | 0.43 |
| 11. | Domestic activity (C) | 0-4 | 0.76 (1.09) | 0.21 | 0.52 | (-0.31) | | | 0.71 |
| 12. | Vocational activity(C) | 0-4 | 0.44 (0.88) | 0.32 | 0.49 | | | | 0.69 |
| 13. | Cognitive (D) | 0-4 | 0.93 (1.28) | 0.23 | 0.37 | | 0.55 | | |
| 14. | Behavioural (D) | 0-4 | 0.31 (0.86) | 0.34 | 0.51 | | 0.83 | | |
| 15. | Emotional/Mood (D) | 0-4 | 0.60 (0.99) | 0.39 | 0.52 | | 0.70 | | |
| 16. | Family support (D) | 0-4 | 0.48 (0.96) | 0.47 | 0.61 | | 0.56 | | |
| 17. | Emotional – staff (D) | 0-4 | 0.54 (0.96) | 0.49 | 0.65 | | 0.75 | | |
| 18. | Discharge (E) | 0-4 | 1.02 (1.27) | 0.27 | 0.57 | | | 0.68 | |
| 19. | Benefits (E) | 0-4 | 0.33 (0.77) | 0.33 | 0.42 | | | 0.49 | |
| 20. | Equipment (E) | 0-4 | 0.48 (0.91) | 0.22 | 0.49 | | | 0.76 | |
| 21. | Community (E) | 0-4 | 0.43 (0.99) | 0.15 | 0.49 | | | 0.61 | |
| 22. | Key working (E) | 0-4 | 1.07 (1.09) | 0.39 | 0.42 | | | 0.42 | |
| 23. | Meetings (F) | 0-2 | 0.83 (0.91) | 0.33 | | | | | |
| 24. | Reports (F) | 0-2 | 0.69 (0.83) | 0.09 | | | | | |
| 25. | Groups (F) | 0-2 | 0.49 (0.76) | 0.23 | | | | | |
| 26. | Clinical attendance (F) | 0-2 | 0.12 (0.44) | 0.15 | | | | | |

*NPTDA: Northwick Park Therapy Dependency Assessment; Extraction Method: Principal Component Analysis. Rotation Method: Promax; all loadings<0.3 removed for clarity. Domains: (A) physical/handling programme, (B) basic functions, (C) activities of daily living, (D) cognitive/ psychosocial /family support, (E) preparing for discharge, (F) additional activities; (G) special input not shown.*

For all 22 NPTDA items included in the exploratory analysis, coefficient-$\alpha$ was 0.74. The corresponding values for the four factors were 0.71 for *Physical*, 0.71 for *Psychosocial*, 0.65 for *Discharge planning* and 0.48 for *Activities*. The correlations between factors ranged from 0.33 (*Activities* and *Psychosocial*) to -0.07 (*Physical* and *Discharge planning*).

Confirmatory factor analysis was conducted on the second randomly selected sample (N=1882) using the four-factor hypothesised model (*Table 30*). For the final model the RMSEA was 0.069, CFI/TLI 0.739/0.701 and the GFI was 0.909, representing a reasonable fit. The final model supported the four-factor hypothesised structure of the NPTDA scale. In summary therefore the fit was with the original domain structure of the scale was remarkably good.

**Table 30: Confirmatory factor analysis: factor loadings for the final model**

| | Item (domain)* | Factor 1 Physical | Factor 2 Psychosocial | Factor 3 Discharge planning | Factor 4 Activities |
|---|---|---|---|---|---|
| **1.** | **Medical (A)** | 0.52 | | | |
| **2.** | **Splinting (A)** | 0.29 | | | |
| **3.** | **Seating (A)** | 0.47 | | | |
| **4.** | **Physical therapy (A)** | 0.41 | | | |
| **5.** | **Tracheostomy (B)** | 0.45 | | | |
| **6.** | **Swallowing (B)** | 0.55 | | | |
| **7.** | **Nutrition (B)** | 0.59 | | | |
| **8.** | **Communication (B)** | 0.44 | | | |
| **9.** | **Speech/language (B)** | 0.45 | | | -0.19 |
| **10.** | **Personal/ self-care (C)** | | | | 0.26 |
| **11.** | **Domestic activity (C)** | | | | 0.68 |
| **12.** | **Vocational activity (C)** | | | | 0.59 |
| **13.** | **Cognitive (D)** | | 0.44 | | |
| **14.** | **Behavioural (D)** | | 0.62 | | |
| **15.** | **Emotional/Mood (D)** | | 0.59 | | |
| **16.** | **Family support (D)** | | 0.52 | | |
| **17.** | **Emotional – staff (D)** | | 0.68 | | |
| **18.** | **Discharge (E)** | | | 0.63 | |
| **19.** | **Benefits (E)** | | | 0.51 | |
| **20.** | **Equipment (E)** | | | 0.50 | |
| **21.** | **Community (E)** | | | 0.44 | |
| **22.** | **Key working (E)** | | | 0.51 | |

**Domains: (A) physical/handling programme, (B) basic functions, (C) activities of daily living, (D) cognitive/ psychosocial /family support, (E) preparing for discharge*

We then examined the change in NPTDA scores between admission and discharge. The results are shown in *Table 31*.

**Table 31: Descriptive statistics of the 26 Northwick Park Therapy Dependency Assessment (NPTDA) items: scores on admission and discharge, and change scores**

| Domain/ Item - Factor | Admission Mean (SD) | Discharge Mean (SD) | Mean difference | 95% CI | Paired t tests t | df | p value |
|---|---|---|---|---|---|---|---|
| **Direct patient care** | | | | | | | |
| **(A) Physical/handling programme (range 0-16)** | | | | | | | |
| 1. Medical (A) | 1.7 (1.1) | 1.2 (1.1) | 0.5 | 0.45, 0.6 | 17.4 | 1417 | <0.0001 |
| 2. Splinting (A) | 0.7 (1.3) | 0.6 (1.2) | 0.1 | -0.01, 0.1 | 1.8 | 1365 | 0.066 |
| 3. Seating (A) | 1.3 (1.3) | 0.6 (1.0) | 0.6 | 0.58, 0.7 | 19.2 | 1365 | <0.0001 |
| 4. Physical therapy (A) | 2.9 (1.1) | 2.2 (1.4) | 0.7 | 0.65, 0.8 | 19.5 | 1365 | <0.0001 |
| **(B) Basic functions (range 0-20)** | | | | | | | |
| 5. Tracheostomy (B) | 0.2 (0.7) | 0.1 (0.4) | 0.1 | 0.07, 0.1 | 6.1 | 1365 | <0.0001 |
| 6. Swallowing (B) | 0.4 (0.9) | 0.2 (0.6) | 0.3 | 0.21, 0.3 | 11.0 | 1416 | <0.0001 |
| 7. Nutrition (B) | 0.6 (1.0) | 0.4 (0.8) | 0.2 | 0.15, 0.2 | 8.8 | 1416 | <0.0001 |
| 8. Communication (B) | 0.4 (1.0) | 0.3 (0.8) | 0.2 | 0.11, 0.2 | 7.1 | 1416 | <0.0001 |
| 9. Speech/language (B) | 0.8 (1.1) | 0.6 (1.0) | 0.3 | 0.23, 0.3 | 10.3 | 1416 | <0.0001 |
| ***Physical (A+B)*** | ***8.2 (5.8)*** | ***5.5 (4.4)*** | ***2.7*** | ***2.5, 3.0*** | ***20.5*** | ***1365*** | ***<0.0001*** |
| **(C) Activities of daily living (range 0-12)** | | | | | | | |
| 10. Personal/ self-care (C) | 1.4 (1.3) | 0.8 (1.1) | 0.5 | 0.4, 0.6 | 13.8 | 1416 | <0.0001 |
| 11. Domestic activity (C) | 0.7 (1.1) | 0.9 (1.1) | -0.1 | -0.2, -0.04 | -3.2 | 1416 | 0.001 |
| 12. Vocational activity(C) | 0.3 (0.8) | 0.5 (0.9) | -0.2 | -0.2, -0.1 | -6.7 | 1416 | <0.0001 |
| ***Activities (C)*** | ***2.4 (2.2)*** | ***2.2 (2.2)*** | ***0.2*** | ***0.1, 0.3*** | ***3.4*** | ***1416*** | ***0.001*** |
| **(D) Cognitive/ psychosocial /family support (range 0-20)** | | | | | | | |
| 13. Cognitive (D) | 1.1 (1.3) | 0.9 (1.2) | 0.3 | 0.2, 0.3 | 6.9 | 1416 | <0.0001 |
| 14. Behavioural (D) | 0.4 (0.9) | 0.3 (0.8) | 0.1 | 0.03, 0.1 | 3.1 | 1416 | 0.002 |
| 15. Emotional/Mood (D) | 0.6 (1.0) | 0.6 (1.0) | -0.01 | -0.06, 0.05 | -0.2 | 1416 | 0.83 |
| 16. Family support (D) | 0.4 (0.9) | 0.5 (1.0) | -0.1 | -0.2, -0.05 | -3.9 | 1416 | <0.0001 |
| 17. Emotional – staff (D) | 0.5 (0.9) | 0.5 (1.0) | -0.03 | -0.08, 0.02 | -1.3 | 1415 | 0.18 |
| ***Psychosocial (D)*** | ***3.0 (3.5)*** | ***2.8 (3.3)*** | ***0.2*** | ***0.02, 0.3*** | ***2.2*** | ***1415*** | ***0.028*** |
| **(E) Preparing for discharge (range 0-20)** | | | | | | | |
| 18. Discharge (E) | 0.6 (1.0) | 1.5 (1.3) | -0.9 | -1.0, -0.9 | -24.4 | 1416 | <0.0001 |
| 19. Benefits (E) | 0.2 (0.6) | 0.5 (0.9) | -0.3 | -0.4, -0.3 | -12.4 | 1416 | <0.0001 |
| 20. Equipment (E) | 0.2 (0.7) | 0.7 (1.1) | -0.5 | -0.6, -0.4 | -16.8 | 1416 | <0.0001 |
| 21. Community (E) | 0.2 (0.7) | 0.7 (1.2) | -0.5 | -0.6, -0.4 | -13.3 | 1416 | <0.0001 |
| 22. Key working (E) | 0.9 (1.0) | 1.2 (1.2) | -0.3 | -0.4, -0.3 | -11.5 | 1416 | <0.0001 |
| ***Discharge planning (E)*** | ***2.09 (2.4)*** | ***4.6 (3.6)*** | ***-2.5*** | ***-2.2, -2.4*** | ***-26.5*** | ***1416*** | ***<0.0001*** |
| **Indirect patient care** | | | | | | | |
| **F) Additional activities (range 0-8)** | | | | | | | |
| 23. Meetings (F) | 0.8 (0.9) | 0.9 (0.9) | -0.1 | -0.2, -0.04 | -3.5 | 1416 | 0.001 |
| 24. Reports (F) | 0.3 (0.6) | 1.1 (0.8) | -0.8 | -0.8, -0.7 | -30.5 | 1416 | <0.0001 |
| 25. Groups (F) | 0.4 (0.7) | 0.7 (0.8) | -0.3 | -0.3, -0.3 | -13.5 | 1416 | <0.0001 |
| 26. Clinical attendance (F) | 0.1 (0.4) | 0.6 (0.5) | -0.03 | -1.0, -0.01 | -2.1 | 1416 | 0.038 |
| ***Total (F)*** | ***1.6 (1.5)*** | ***2.8 (1.8)*** | ***-1.2*** | ***-1.3, -1.1*** | ***-2419*** | ***1416*** | ***<0.0001*** |

*items for which therapy inputs increase from admission to discharge are shown in red*

The first three factors showed a significant reduction overall i.e., (A+B) *Physical* ($p<0.0001$), (C) *Activities* ($p=0.001$) and (D) *Psychosocial* ($p=0.028$). However, within these three subscales, certain individual items increased towards discharge – namely Vocational activity (leisure) ($p<0.0001$), Domestic (community) activity ($p=0.001$) and Family support ($p<0.0001$). The remaining factor (E) *Discharge planning* and the indirect patient care - (F) additional activities, both show a significant increase in their scores ($p <0.0001$) between admission and discharge. These findings resonate with clinical experience that the focus of a rehabilitation goals changes between the early and late stages of the programme; the early stages focussing on re-acquisition of basic skills and the latter focussing more on plans for discharge and function in the community.

*Table 32* summarises the relationship of the NPTDA with rehabilitation complexity and functional independence on admission. As expected, the admission scores for the NPTDA scale were significantly correlated with both the RCS scores and FIM scores although the correlation was modest. The negative correlation with the FIM scale, particularly with the physical part of the NPTDA scale confirms that patients who are more dependent generally have higher therapy requirements, which is expected. However, weaker correlations with the other subscales confirms that the NPTDA measures aspects of therapy requirements that are not well reflected by the FIM score.

**Table 32: Pearson correlations between NPTDA direct patient care with RCS and FIM based on total and subscales scores on admission**

| | NPTDA Direct patient care (A+B+C+D+E) | Subscales | | | |
|---|---|---|---|---|---|
| | | Physical (A+B) | Activities (C) | Psychosocial (D) | Discharge planning (E) |
| **RCS** | 0.30** | 0.44** | -0.15** | 0.17** | -0.03 |
| **Care + Medical + Nursing** | 0.23** | 0.36** | -0.20** | 0.15** | -0.07** |
| **Therapy** | 0.29** | 0.35** | 0.01 | 0.13** | 0.07** |
| **FIM** | -0.25** | -0.49** | 0.28** | -0.02 | 0.15** |
| **Cognitive** | -0.25** | -0.36** | 0.17** | -0.22** | 0.11** |
| **Motor** | -0.21** | -0.46** | 0.27** | 0.06* | 0.13** |

*Domains: (A) physical/handling programme, (B) basic functions, (C) activities of daily living, (D) cognitive/ psychosocial /family support, (E) preparing for discharge; RCS: Rehabilitation Complexity Scale; FIM: Functional Independence Measure ; *p<0.05; **p<0.01*

In summary, this evaluation confirmed a four-factor structure of the NPTDA although Cronbach's coefficient-$\alpha$ for the whole scale, as well as for the restricted scale (22 items), was within the acceptable range. The relationships between the NPTDA scale and the RCS and FIM scales confirmed our expectations and further supported the concurrent validity of the scale. However, as expected the correlations were relatively low, suggesting that the scales indeed measure different aspects of patient rehabilitation inputs.

### Limitations of the study

- Although the sample size exceeded the usual standards for Factor analysis, these analyses were carried out on two sub-samples that were not independent.
- NPTDA scores were only available from a proportion of the sample which could have led to a degree of selection bias.

Despite the acknowledged limitations, the NPTDA has shown acceptable internal reliability, good construct and concurrent validity for measuring multidisciplinary therapy interventions in neurorehabilitation. It is responsive to change during neurorehabilitation between admission and discharge in a manner that resonates with clinical experience. The findings suggest that the NPTDA scale is a rehabilitation tool that provides useful and reliable estimates of multidisciplinary therapy interventions in patients with complex disability undergoing treatment in specialist neurorehabilitation settings.

# Adapted versions of the NPTDA for different settings

Specialist neurorehabilitation services are not homogeneous. For example, some services cater specifically for patients with severe physical disabilities, whilst others provide for more mobile patients with predominantly cognitive/behavioural problems. Whilst some therapy interventions will be common to all groups, other interventions will vary depending on the type of caseload.

The NPTDA described above was developed primarily for patients with physical and mixed neurological disabilities. Feedback during the early stages of rollout indicated that some adaptation was required for two key specialist rehabilitation service areas which were (a) cognitive behavioural and (b) paediatric rehabilitation.

### Approach to adaptation

As with our adaptations of other tools, our approach was to maintain the integrity, structure and scoring systems of the base instrument so far as possible, but to identify items that were redundant in these different specialist service areas, and replace them with items that were more relevant to their type of service.

The process involved the following stages:

1. In the first instance the teams applied the original NPTDA in a small number of patients.
2. One or more workshops / focus groups were then held with the team to discuss any the strengths and weakness of the tool in their setting, to identify items that were less relevant and discuss what should replace them.
3. A revised tool was then drafted and piloted.
4. The tool was refined through an iterative process of feedback and revision, until it stabilised.

### Adaptation for cognitive behavioural services

We worked with Dr Simon Fleminger (consultant neuropsychiatrist) and teams in the three specialist Level 1 rehabilitation services in London that delivered primarily cognitive behavioural rehabilitation (The Lishman Unit, Mausdley Hospital; the Brain Injury Rehabilitation Unit (BIRU) in Edgware; and the Blackheath Brain Injury Rehabilitation Centre). Three workshops were held and iterative development took place over the course of six months.

**Key changes were:**

- Four items in the physical/handling section (related to splinting, and seating which were largely irrelevant in this group) were replaced with four items relating to psychiatric and risk management, including mental capacity and legal issues and management under section of the Mental Health Act.
- One item (Tracheostomy management) in the basic functions domain was replaced with a single item encompassing general physical management.
- The option to record management of a serious untoward incident (e.g. patient absconding etc.) as a multidisciplinary activity under 'Additional activities'

### Adaptation for children's services

We worked with the two signposted children's specialist rehabilitation services (The Children's Trust, Tadworth Court and Chailey Heritage Brain Injury Rehabilitation Centre) to adapt the NPTDA for use in this context. A total of five workshops were held between the two centres. After much debate and discussion, and extensive trialling of various versions, we established that no change was required to the overall structure of the tool, but some rewording of items was required for the paediatric context. For example, 'vocational/leisure/computers' was re-titled to include 'play therapy and education'. For a Table showing the comparative contents of the versions, see *Appendix 4.9.*

Data collection with these tools is well underway. The Children's version is structurally so similar to the adult version as to make further evaluation of its construct validity unnecessary. Construct validity of the Cognitive behavioural version will be evaluated in future once the dataset reaches sufficient size, but as there are relatively few episodes admitted to these services, the numbers accumulated are as yet are insufficient for factor analysis.

*Table: 33* summarises the findings from psychometric evaluation of the NPTDA according to the Medical Outcomes Trust framework

**Table 33: Psychometric evaluation of the Northwick Park Therapy Dependency Scale according to the Medical Outcomes Trust framework**

| Attribute and criteria | Evaluation |
|---|---|
| **Conceptual and measurement model:** ***The rationale for and description of the concept and the populations that the measure is intended to assess*** | |
| **Clinical content and design** | A 30-item ordinal rating scale to quantify an individual's needs for therapy interventions in a neurorehabilitation setting.<br>Subscales: 5 domains of direct 'hands on' intervention - (Physical handing (4 items), Basic functions (5 items), Activities of daily living (3 items) , Cognitive/Psychosocial support (5 items), and discharge panning (5 items) - all rated on a scale 0-4;<br>One subscale of Additional activities relating to indirect patient care (4 items on a scale of 0-2 ). Total score range 0-100<br>Translates by a computerised algorithm into an assessment of therapy staff time separated by discipline.[114] |
| **Dimensionality** | Principal component analysis demonstrates 4 factors in the direct intervention subscale exactly reflecting the domain structure, except that Physical handing and Basic functions map onto a single factor |
| **Reliability:** ***The degree to which the instrument is free from random error*** | |
| **Internal consistency and homogeneity** | Cronbach's alpha: For the 22-item direct NPTDA (alpha = 0.74); Corresponding values for the four factors are *Physical*, $\alpha$=0.71 for *Psychosocial* $\alpha$=0.71, *Discharge planning* $\alpha$=0..65 and 0.48 for *Activities* $\alpha$=0.48. Heterogeneity was expected for the full 26 item NPTDA scale. Item-total correlations ranged from 0.15 to 0.49, and less than 50% being above 0.30.<br>Item-total correlations: BCN section: 0.34 – 0.88; SNN section 0.18 – 0.31 (also reflects lower positive score rates for these items). |
| **Reproducibility** | Inter-rater reliability: Not formally tested |
| **Validity:** ***The degree to which the instrument measures what it purports to measure*** | |
| **Content** | Within developmental design – based on clinicians' expert opinion, and several rounds of observed activity analysis and work sampling in a hospital setting.[114] |
| **Criterion-related** | Not testable - no accepted gold standard currently exists. |
| **Construct** | One study shows moderate but significant positive correlation with the simultaneously recorded Rehabilitation Complexity Scale (rho 0.17-0.44, p<0.01) and negative correlations with the Functional Independence Measure (0.15-0.49, 0<0.01), especially in the physical domains |
| **Discrimination and responsiveness:** ***Ability to detect change over time where real changes occur*** | |

| | |
|---|---|
| **Change: admission to discharge** | Responsive to change between admission and discharge with overall reduction in therapy interventions, but increase in individual items (e.g. vocational rehabilitation and discharge planning which resonate with clinical experience |
| **Interpretability:** ***Whether easily understood meaning can be assigned to quantitative scores*** | |
| **Clinical meaning** | Teams report general satisfaction with the content being representative of their inputs |
| **Burden:** ***The time, effort or other demands of administering the instrument*** | |
| **Time to administer** | The mean time to complete the NPTDA for an experienced user was approximately 6-7 minutes.[114] |
| **Alternative modes of administration** | None |
| **Cultural and language adaptations** | Adapted versions exist for cognitive behavioural rehabilitation and childrens' rehabilitation settings |

# Outcomes

## *The Neurological Impairment Scale*

As noted above, rehabilitation often occurs against a background of changing neurological impairment, which may be either improvement or deterioration depending upon the underlying condition. It is therefore necessary to have a standardized assessment of neurological impairment, against which any change in functional independence can be evaluated. A wide range of impairment sets exist for specific conditions. Well-known examples include the NIH scale for stroke,[116] the ASIA scale for spinal cord injury[117] or the EDSS for multiple sclerosis.[118] However these are not necessarily applicable across the broader range of neurological conditions.

The Neurological Impairment Scale (NIS) evolved from the impairment set that was developed alongside the UK FIM+FAM in the 1990s (see below), and has been included within the original 'minimum dataset' of the UK FIM+FAM software programme since 1999. The original checklist provided a crude identification of different types of impairment, but gave no indication of severity and was therefore insensitive to change. Consequently, later versions of the database not only expanded the range of impairments but also introduced a simple grading of severity ('None', 'Mild', 'Moderate', 'Severe'). After the WHO International Classification of Functioning (ICF) was published in 2002,[119] the NIS items were mapped onto ICF codes to support data collection using the common language of the ICF. Over the course of some 15 years and several iterations, the tool has evolved into an ordinal measure of impairment that has potential applicability across a wide range of neurological conditions. The current version of the NIS (version 8), comprises 17 items (each rated 0-2 or 0-3 giving a total score range 0-50). See *Appendix 4.11 ?4.9.*

At the onset of this programme, however, there had been no formal evaluation of its internal psychometric properties, nor of its usefulness as a predictor of disability or of potential to make functional gains in rehabilitation. Neither was it clear when and by whom the NIS should be recorded in the rehabilitation process. In the course of this programme we performed the first examination of the reliability and validity of the NIS as a measure of an individual's specific and overall neurological impairment in a sample of patients with a diverse range of neurological disabilities. The full paper has been published in Disability and Rehabilitation[120], but an abbreviated version is presented here.

The objectives of this evaluation were to:

a) Investigate its construct validity through an exploratory factor analysis and examination of internal consistency.

b) Determine inter-rater reliability and to compare ratings recorded by the medical staff and the multidisciplinary team..
c) Examine concurrent validity in terms of the relationship between changes in impairment and disability.
d) Assess predictive validity of the NIS as a predictor of outcome and functional gains made during inpatient rehabilitation.

## *Design.*

A cohort analysis of patients admitted for in-patient neurorehabilitation.

In Part 1, we used a principal components analysis with Varimax rotation to examine the dimensionality and internal consistency of the NIS in a multicentre sample (n=428) from nine specialist neurorehabilitation units, as part of a parallel NIHR-funded study within our research group.[121] The NIS was rated by the multidisciplinary (MD) team at the point of discharge into the community.

In Part 2, we performed a more detailed analysis of inter-rater reliability and the relationship between NIS and functional outcome (UK FIM+FAM scores) in the cohort sample (n=94) from just one of the contributing centres under this programme.

This unit is the lead centre of training and dissemination of the UK FIM+FAM and NIS, so use of both scales was well established in this setting and all staff were fully trained in their use. Both scores were recorded as part of routine practice by the multidisciplinary team within 10 days of admission (n=94), and repeated within 7 days of discharge (n=73) by the end of the study period**.** In addition, the medical team independently undertook NIS ratings, based on presentation of findings collected at the admission medical clerking within 48 hours of admission for n=77 patients.

To test inter-rater reliability of this method, two doctors independently recorded NIS scores for a sub-sample of n=47 patients. Patients were included in the analysis sample if they had (a) paired NIS ratings at admission and discharge (n=73) or (b) paired ratings by the medical and MD team (n=77) (some patients were included on both counts). The demographics for these two patient samples are shown in *Table 34*.

**Table 34: Demographics of the study populations for Parts 1 and 2**

| | Part 1 n=428 | | Part 2 n=94 | |
|---|---|---|---|---|
| | Mean | SD | Mean | SD |
| **Demographics** | | | | |
| **Age (years)** | 49.5 | 15.3 | 42.9 | 14.5 |
| **Length of stay (days)** | - | - | 93 | 62 |
| **Mean time since onset (days)** | - | - | 104 | 71 |
| | n | % | n | % |
| **Males: Females** | 270:158 | 63:37 | 59:35 | 63:37 |
| **Diagnosis** | | | | |
| **Acquired brain injury** | 315 | 74% | 79 | 85% |
| **Vascular (stroke, SAH)** | 212 | 50% | 41 | 44% |
| **Traumatic** | 63 | 15% | 28 | 30% |
| **Other (e.g. Hypoxic/inflammatory)** | 40 | 9% | 10 | 11% |
| **Spinal cord injury** | 38 | 9% | 7 | 7% |
| **Guillain-Barré and other peripheral neuropathies** | 26 | 6% | 5 | 5% |
| **Multiple sclerosis** | 21 | 5% | 7 | - |
| **Others** | 27 | 6% | 3 | 3% |

*SAH = Sub-arachnoid Haemorrhage*

## *Part 1: Principal components analysis with Varimax rotation*

*Table 35* presents the results of the principal components analysis after Varimax rotation. In fact the pattern of loadings on the first two unrotated principal components and on the two rotated factors were almost identical. Two components had eigenvalues >2, which together accounted for 35 % of the variance in total scores.

The most significant source of variance is explained by a *Physical* component and the second major source of variance is a *Cognitive* component. Only two of the 15 items ('Hearing' and 'Other') achieved loadings of <0.4 on both components mainly due to a preponderance of zero scores (90% and 85% respectively).

In subsequent analyses, 'Hearing' was included in the cognitive subscale. The 'Other' score (which most commonly includes impairments such as seizures or pressure sores) was included in the physical subscale. Cronbach's alpha for the full 17-item NIS was 0.75. For the ten Physical and seven Cognitive items, Cronbach's alphas were 0.76 and 0.67 respectively.

**Table 35: NIS score ranges, item-total correlations and loadings on first two principal components using Varimax rotation**

| | Descriptives | | | Rotated component matrix | |
|---|---|---|---|---|---|
| | Median (IQR) | Range | Item total Correlations | Factor 1 Physical | Factor 2 Cognitive |
| eigenvalue | | | | 3.6 | 2.3 |
| % variance | | | | 20.4 | 14.6 |
| **NIS Items** | | | | | |
| Left upper limb | 0 (0-1) | 0-3 | 0.43** | **0.49** | |
| Right upper limb | 1 (0-2) | 0-3 | 0.44** | **0.50** | |
| Left lower limb | 1 (0-2) | 0-3 | 0.47** | **0.68** | |
| Right lower limb | 1 (0-2) | 0-3 | 0.50** | **0.69** | |
| Trunk | 1 (0-1) | 0-2 | 0.50** | **0.67** | |
| Tone /contractures | 1 (0-2) | 0-3 | 0.63** | **0.70** | |
| Sensation | 1 (0-2) | 0-3 | 0.55** | **0.63** | |
| Perception | 0 (0-1) | 0-3 | 0.38** | | **0.50** |
| Speech and language | 1 (0-2) | 0-3 | 0.45** | | **0.58** |
| Cognitive | 1 (1-2) | 0-3 | 0.33** | | **0.81** |
| Behaviour | 0 (0-0) | 0-3 | 0.23** | | **0.64** |
| Mood | 1 (0-1) | 0-3 | 0.40** | | **0.53** |
| Vision | 1 (0-1) | 0-3 | 0.31** | | **0.56** |
| Hearing | 0 (0-0) | 0-2 | 0.09 | | |
| Pain | 1 (0-1) | 0-3 | 0.50** | **0.51** | |
| Fatigue | 1 (1-2) | 0-3 | 0.52** | **0.54** | |
| Other | 0 (0-0) | 0-2 | 0.12* | | |

*** significant at p<0.001, * significant at p<0.05*

### Part 2: Reliability and relationship with functional outcome

#### *a) Inter-rater reliability*

*Table 36* shows the level of inter-rater agreement found between the different ratings. Within the medical team, there was high overall agreement between the two doctors reflected by a kappa coefficient of 0.81 for total NIS score and ICC 0.95 (95% confidence interval 0.91-0.97). Item-by-item agreement ranged from 'substantial' to 'almost perfect' with the exception of the Fatigue item.

**Table 36: Inter-rater agreement: Item-by-item linear-weighted Kappa coefficients interpreted according to Landis and Koch 1977**

| | Agreement between two doctors (n=47) | | | Agreement between Medical and MD Team scores (n=94) | | |
|---|---|---|---|---|---|---|
| **Item** | **Kappa** | **95% CI*** | **Interpretation** | **Kappa** | **95% CI** | **Interpretation** |
| Left upper limb | 0.94 | 0.72-1.0 | Almost perfect | 0.76 | 0.61-0.92 | Substantial |
| Right upper limb | 0.90 | 0.69-1.0 | Almost perfect | 0.74 | 0.59-0.90 | Substantial |
| Left lower limb | 0.93 | 0.72-1.0 | Almost perfect | 0.82 | 0.66-0.97 | Almost perfect |
| Right lower limb | 0.85 | 0.65-1.0 | Almost perfect | 0.76 | 0.92-0.59 | Substantial |
| Trunk | 0.87 | 0.64-1.0 | Almost perfect | 0.43 | 0.27-0.59 | Moderate |
| Tone | 0.63 | 0.44-0.82 | Substantial | 0.45 | 0.31-0.60 | Moderate |
| Sensation | 0.82 | 0.59-1.0 | Almost perfect | 0.65 | 0.51-0.79 | Substantial |
| Perception | 0.85 | 0.60-1.0 | Almost perfect | 0.63 | 0.47-0.79 | Substantial |
| Speech | 0.90 | 0.68-1.0 | Almost perfect | 0.81 | 0.66-0.96 | Almost perfect |
| Cognitive | 0.84 | 0.63-1.0 | Almost perfect | 0.66 | 0.52-0.79 | Substantial |
| Behaviour | 0.94 | 0.69-1.0 | Almost perfect | 0.13 | 0-0.27 | Slight |
| Mood | 0.75 | 0.54-0.96 | Substantial | 0.57 | 0.43-0.72 | Moderate |
| Vision | 0.80 | 0.56-1.0 | Substantial | 0.68 | 0.52-0.84 | Substantial |
| Hearing | 0.79 | 0.52-1.0 | Substantial | 0.77 | 0.59-0.95 | Substantial |
| Pain | 0.76 | 0.58-0.94 | Substantial | 0.43 | 0.29-0.56 | Moderate |
| Fatigue | 0.58 | 0.40-0.77 | Moderate | 0.39 | 0.27-0.52 | Fair |
| Other | 0.84 | 0.63-1.0 | Almost perfect | 0.44 | 0.29-0.59 | Moderate |
| **Total NIS Kappa** | **0.81** | **0.63-0.99** | **Almost perfect** | **0.69** | **0.56-0.95** | **Substantial** |
| **ICC** | **0.95** | **0.91-0.97** | | **0.92** | **0.88-0.95** | |

**95% Confidence intervals (CI) were calculated as +/-1.96s Standard Error and the upper limit truncated at a maximum of 1.0*

Less strong agreement was expected between the ratings by the medical and MD team, as up to 10 days elapsed between the assessments. Nevertheless, agreement in total scores was still acceptable (Kappa 0.69 for total NIS score, ICC 0.92 (95%CI 0.88-0.95). Agreement for individual items was moderate to strong for 10/13 items. Only slight or fair agreement was observed, however, for the items for *Behaviour* and *Fatigue* (see discussion).

Medical teams tended to record slightly lower ratings than the MD team, although this did not reach significance, either for the total scores or at item level. The MD team ratings were considered to be more reliable and were therefore used in the further evaluation of NIS as a predictor of functional outcome.

***b) Concurrent validity – the relationship between impairment and disability.***

*Table 37* summarises the admission, discharge and change scores for the NIS and UK FIM+FAM scores. Both the NIS and the FIM+FAM showed significant changes between admission and discharge for total scores and subscales.

**Table 37: Descriptive statistics for NIS and UK FIM+FAM scores as rated by the MD Team on admission and discharge (n=73)**

| Neurological Impairment Scale (NIS) | | | | | | Paired sample t tests | | |
|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Median | IQR | Range | t | p | ES* |
| **Admission** | | | | | | | | |
| Physical Subscale | 14.0 | 6.7 | 13 | 10 - 19 | 1 - 28 | - | - | - |
| Cognitive Subscale | 8.2 | 6.2 | 7 | 4 - 11 | 0 - 21 | - | - | - |
| Total Score | 22.2 | 11.6 | 19 | 15 - 28 | 4 - 48 | - | - | - |
| **Discharge** | | | | | | | | |
| Physical Subscale | 11.8 | 6.8 | 11 | 7 - 17 | 0 - 27 | - | - | - |
| Cognitive Subscale | 6.7 | 5.6 | 5 | 2 - 8 | 0 - 21 | - | - | - |
| Total Score | 18.3 | 11.5 | 15 | 10 - 22 | 3 - 46 | - | - | - |
| **Change** | | | | | | | | |
| Physical Subscale | -2.2 | 2.9 | -2 | -4 - 0 | -11 - 5 | 6.4 | <0.001 | 0.73 |
| Cognitive Subscale | -1.5 | 2.4 | -1 | -3 - 0 | -11 - 3 | 5.3 | <0.001 | 0.64 |
| Total Score | -3.7 | 4.1 | -3 | -7 - -1 | -18 - 5 | 7.8 | <0.001 | 0.90 |
| **UK Functional Assessment Measure (FIM+FAM)** | | | | | | | | |
| | Mean | SD | Median | IQR | Range | t | p | ES* |
| **Admission** | | | | | | | | |
| Motor Subscale | 53.9 | 30.0 | 51 | 21 - 86 | 16 - 108 | - | - | - |
| Cognitive Subscale | 56.3 | 27.0 | 61 | 32 - 79 | 14 - 95 | - | - | - |
| Total Score | 110.1 | 52.0 | 123 | 63 - 156 | 30 - 189 | - | - | - |
| **Discharge** | | | | | | | | |
| Motor Subscale | 74.3 | 33.3 | 87 | 47 - 102 | 16 - 111 | - | - | - |
| Cognitive Subscale | 69.5 | 24.6 | 77 | 59 - 88 | 14 - 98 | - | - | - |
| Total Score | 143.8 | 55.6 | 166 | 109 - 188 | 30 - 206 | - | - | - |
| **Change** | | | | | | | | |
| Motor Subscale | 20.4 | 16.5 | 17 | 6 - 33 | -2 - 62 | -10.5 | <0.001 | -1.26 |
| Cognitive Subscale | 13.3 | 11.9 | 11 | 5 - 18 | -5 - 55 | -9.5 | <0.001 | -1.14 |
| Total Score | 33.7 | 23.4 | 31 | 14 - 52 | -3 - 93 | -12.3 | <0.001 | -1.48 |

**Effect size (ES) calculated as Cohen's d allowing for the correlation between the mean*

*Table 38* shows the correlation between admission and change scores. Strong negative correlations were seen between NIS Physical and FIM+FAM Motor subscales on admission (Pearson r = -0.86), and similarly between the respective cognitive subscales (r= -0.90). Weaker, but still significant correlations were seen between physical and cognitive domains of the respective scales (r 0.62, p<0.001 in each case).

Changes in NIS subscale scores were significantly correlated with change in their respective components of the UK FIM+FAM (r=-0.51 to -0.56). However, although there was a strong negative correlation between change in FIM+FAM Motor score and the NIS cognitive score on admission (r= -0.45, p<0.001), no such relationship was seen with the admission NIS physical score.

**Table 38: Pearson correlations for admission and change scores (n=73)**

| | Admission scores | | | | | | Change scores | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | NIS | | | FIM+FAM | | | NIS | | | FIM+FAM | |
| | **Phys** | **Cog** | **Total** | **Motor** | **Cog** | **Total** | **Phys** | **Cog** | **Total** | **Motor** | **Cog** |
| **Admission** | | | | | | | | | | | |
| **NIS Physical** | | | | | | | | | | | |
| **NIS Cognitive** | **.62*** ** | | | | | | | | | | |
| **NIS Total** | **.91*** ** | **.89*** ** | | | | | | | | | |
| **FIM+FAM Motor** | **-.86*** ** | **-.62*** ** | **-.85*** ** | | | | | | | | |
| **FIM+FAM Cognitive** | **-.62*** ** | **-.90*** ** | **-.83*** ** | **.67*** ** | | | | | | | |
| **FIM+FAM Total** | **-.81*** ** | **-.82*** ** | **-.90*** ** | **.92*** ** | **.90*** ** | | | | | | |
| **Change score** | | | | | | | | | | | |
| **NIS Physical** | -.18 | .24* | .02 | .09 | -.17 | -.20 | | | | | |
| **NIS Cognitive** | -.28* | **-.42*** ** | -.39** | .27* | .35* | .28* | .19 | | | | |
| **NIS Total** | -.29* | -.07 | -.21 | .22 | .09 | .17 | **.82*** ** | **.72*** ** | | | |
| **FIM+FAM Motor** | -.16 | **-.45*** ** | -.34* | -.06 | .36** | .12 | **-.56*** ** | -.02 | **-.42*** ** | | |
| **FIM+FAM Cognitive** | .07 | .17 | .13 | -.22 | -.42*** | -.35** | -.28* | **-.51*** ** | **-.49*** ** | .34** | |
| **FIM+FAM Total** | -.08 | -.23* | -.17 | -.15 | .05 | -.13 | **-.54*** ** | -.27* | **-.54*** ** | **.88*** ** | **.75*** ** |

**p < 0.05, ** p < 0.01, *** p < 0.001*
*NIS=Neurological Impairment Scale, FIM+FAM = UK Functional Assessment Measure*

Within this dataset, approximately two-thirds (n=49) were 'impairment responders' (i.e. their NIS scores reduced by two or more points between admission and discharge. The remaining 24 impairment 'non-responders' showed no such reduction - indeed ten of them showed an increase in total NIS score ranging from 1-5.

- Both groups made significant functional gains during in-patient rehabilitation as illustrated in *Figure 22*.
- However, the 'impairment 'responders' showed a significantly greater change in FIM+FAM motor score (mean change 24.8, sd 17.0) compared with 'non-responders' (mean change 11.4, sd 11.0) giving a mean difference of 13.3 (95%CI 6.6,20.0) t -4.4 p<0.001).
- Similarly, impairment responders showed a greater change in FIM+FAM cognitive score (mean change 16.6, sd 12.6) than the non-responders (mean change 6.5, sd 6.6) giving a mean difference of 10.0 (95% CI 5.5, 14.5) t -4.5 p<0.001).

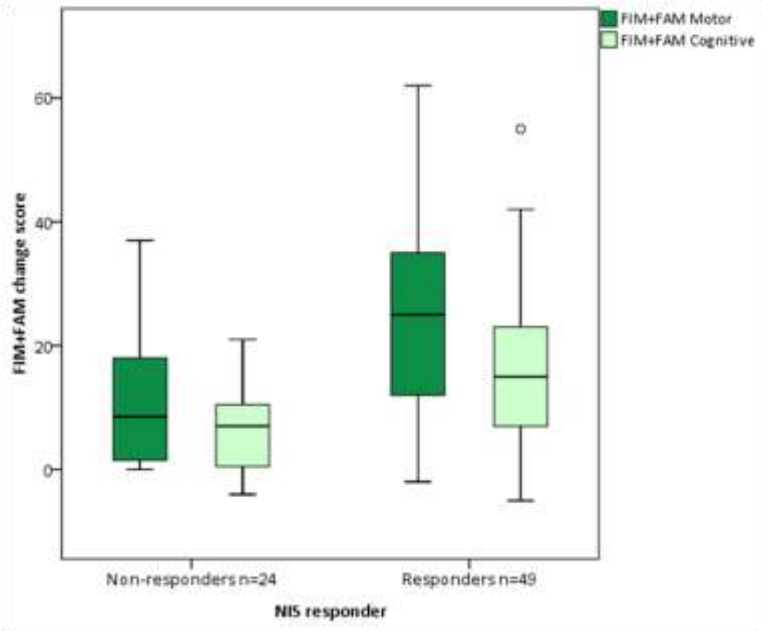



**Figure 22: Box and whiskers plots for change in total FIM+FAM domain scores between admission and discharge in the impairment 'responder' and 'non-responder' groups**

***Legend***
*Figure 22 shows a box and whiskers plots of the FIM+FAM change scores, in patients who did and did not demonstrate change in NIS score during their rehabilitation programme. Both groups improved overall, but impairment 'responders' made significantly greater gains in both motor and cognitive function than the 'non-responders'.*

***c) Predictive validity - the NIS as a predictor of outcome and functional gains made during inpatient rehabilitation.***

To determine whether the NIS adds to the prediction of functional outcome, two linear multiple regression models were tested using total discharge FIM+FAM score as the independent variable.

In both of these, the admission FIM+FAM score was entered first and accounted for 82% of the variance. The NIS contributed a further 5-6%, so that the combination of baseline disability and impairment (baseline or change) predicted 88% of the level of overall disability outcome score.

To examine the factors predictive of *change* in function during rehabilitation, the following variables were entered stepwise into the model using FIM+FAM total change score as the independent variable: a) Admission FIM+FAM total score, b) Admission NIS physical and cognitive scores, c) NIS physical and cognitive change scores. Within this model, admission FIM+FAM score and admission NIS physical and cognitive scores and change in NIS cognitive score were all excluded as predictor variables. Only the change in NIS physical score was entered into the model where it accounted for 29% of the variance.

In summary, this first psychometric examination of the NIS has provided evidence for its scaling properties, reliability and concurrent and predictive validity. Exploratory factor analysis in a large multi-centre sample demonstrated two distinct principal components, which led to the identification a 10-item sub-scale of physical impairment, and a seven-item sub-scale of cognitive impairment each with acceptable internal consistency. Although the 'Hearing' and 'other' (e.g. seizures, pressure sores) items did not load on either factor, this reflected a preponderance of zero scores in this sample. The items have been retained in the scale for their clinical importance.

Inter-rater agreement was very high, both between two doctors and between the doctors and the MD Team, suggesting that the NIS can be applied by either group. However, the MD Team assessment was thought to provide a more comprehensive assessment and was more likely to identify subtle impairments, such as mood, pain, behaviour and fatigue as they emerged over time.

The expected concurrent and divergent relationships were seen between the physical and cognitive domains of the NIS and UKFIM+FAM. However, the moderate correlations between these measures confirms that the underlying constructs of impairment and disability are distinct, each requiring measurement in their own right.

In this series, functional gains during rehabilitation (as measured by change in total FIM+FAM) were significantly predicted both by the level of cognitive impairment on admission, and by change in physical impairment (as measured by the NIS). To our knowledge, this is the first study to examine the influence of changing neurological impairment on functional gains recorded by UK FIM+FAM. Our findings suggest that the NIS can make a useful contribution to the prediction of functional outcome in a mixed-diagnosis group of patients with severe/complex neurological disability.

**Limitations of the study:**

- Both parts of the study were conducted on a highly selected group of patients with severe complex neurological disability, admitted for rehabilitation in Level 1 (tertiary specialist) services in London. Evaluation needs to be expanded in a multi-centre analysis, across a range of clinical settings and with other samples of neuro-rehabilitation patients and clinicians.
- The fact that a tool has been in use for some time does not necessarily mean that it is a good as it can be. As a result of this study, clinical teams highlighted a number of areas of shortfall in the NIS including the evaluation of musculoskeletal impairment and bladder and bowel dysfunction

Nevertheless, our findings showed the NIS to be a promising measure of neurological impairment, suitable for use across a broad range of neurological conditions. They demonstrate that, even in its current form, the NIS can provide useful information for adjustment, over and above the admission FIM+FAM score, as a predictor of functional gain. Such information will assist the interpretation of functional outcomes from inpatient rehabilitation of people with complex neurological disabilities.

### The NIS-Trauma

Towards the end of the programme, the BSRM Trauma working Party (Trauma WP) developed core standards for specialist rehabilitation following trauma[122] in preparation for a national clinical audit. As part of this development the group developed an extension of the NIS to include impairments relevant to the cohort of patients admitted for specialist rehabilitation following multiple trauma (including fractures, limb loss, vascular and visceral injury. However, the additional content has not yet been subject to formal evaluation.

# The UK Functional Independence Measure (FIM) + Functional Assessment Measure (FAM)

## *Background*

Global measures of disability, such as the Functional Independence Measure (FIM)[55] and Functional Assessment Measure (FIM+FAM)[123] are widely used internationally to measure outcome from inpatient rehabilitation programmes. At the individual clinical level, they provide valid and reliable information about a person's requirements for assistance with essential tasks of daily living, and on a group level they can be used to measure and compare outcomes across different practices and populations. Consequently, it is important to understand their metric properties in the population in which they are to be used.

The FIM is an 18-item ordinal measure of disability which includes 13 motor items and 5 cognitive items.[124] It was developed in the 1980s by a national task force in the United States (US) and is now one of the most commonly used generic outcome measures in rehabilitation. Its psychometric properties have been very thoroughly evaluated in the world literature.[124-127] The Functional Assessment Measure does not stand alone (hence the abbreviation "FIM+FAM") but adds a further 12 items to the FIM primarily addressing cognitive and psychosocial function.

The original US version of the FAM was developed in the early 1990s, for evaluating outcomes after traumatic brain injury.[123, 128] Although the US FAM and original training materials are still accessible from the TBI COMBI (Centre for Outcomes Measurement in Brain Injury) website,[129] the US version is no longer actively maintained or centrally collated.

The UK version of the FAM was developed in the mid-1990s by the United Kingdom FIM+FAM Users Group[83] in collaboration with the US originators, to translate it into UK-English and address the known subjectivity and inconsistency of some items. The resulting tool was shown to have improved reliability and utility in comparison with the US version.[83] It has continued to be revised and developed, with the addition of a six-item Extended activities of Daily Living (EADL) module (to extend the upper range of the instrument) and an active programme for training and accrediting users.[130]
The two versions are structurally similar, so that the psychometric performance of the US version has relevance for the UK FIM+FAM. However, they are sufficiently different for the UK version to require validation in its own right.

Although originally conceptualised for use with traumatic or diffuse brain injury, many of the FAM items are more widely applicable in other neurological conditions, including spinal cord injury and progressive neurological conditions. For this reason, the UK FIM+FAM has gained in popularity over the last decade, effectively taking over where development of the US version ceased and it is now the version that continues to be promoted and developed. It has been adopted in the UKROC database as the principal outcome measure for specialist rehabilitation,[72, 131] and is increasingly being explored as an outcome measure for rehabilitation in other countries, including Australia, New Zealand, Europe and South America. Prior to the start of this programme there was a sizeable pool of published papers addressing various psychometric properties of the US and the UK FIM+FAM, but no attempt had been made to assimilate or examine the evidence within the framework of the Medical Outcomes Trust.[63] In addition, much of the existing literature was focussed on its application in the context of traumatic brain injury and it was therefore appropriate to examine its psychometric properties in the broader neurorehabilitation group in which it will be applied. To address these gaps in the literature we published a paper[132] in two parts:
Part 1 presented brief systematic review and assimilation of the existing literature on the psychometric properties of both the original US version and the UK FIM+FAM.

In Part 2, we addressed the identified gaps in the literature, using combination of parametric and non-parametric techniques to explore dimensionality, internal consistency and responsiveness of the UK FIM+FAM in a large consecutive cohort of in-patients representing the diagnostic diversity of a general neurorehabilitation sample. A paper describing this analysis has been publishedin Disability and Rehabilitation[132] so an abbreviated summary is presented here.

## *Part 1 Systematic review*

### Methods

To identify existing studies on psychometric aspects of the US and the UK FIM+FAM we searched the following databases using the search terms *Functional Assessment Measure.mp*, and *FIM+FAM.mp*: Medline 1948 – November 2012, Embase 1980 – November 2012, PsycINFO 1806 – November 2012. Studies concerned with the psychometric properties of the FIM+FAM, as well as studies that were not primarily psychometric but might report relevant statistics (e.g. predictive validity), were identified by two investigators (LTS and RJS) on the basis of the title or abstract.

### Results

We recovered 16 articles reporting on the psychometric qualities of the US FIM+FAM[123, 128, 133-146] and seven[50, 59, 72, 83, 130, 147, 148] on the UK FIM+FAM (six relating to the main scale and one to the EADL module). *Appendix 4.12* summarises the existing literature and also highlights the contribution of new psychometric data presented in this article. The 16 articles on the US version reported a range of important psychometric properties including utility, reliability, validity, dimensionality (i.e. factor structure), responsiveness and floor/ceiling effects. In general the US FIM+FAM had good psychometric properties, although several papers raised concerns about ceiling effects when used in outpatient or community settings.

The seven papers on the UK version reported good psychometric properties for responsiveness, utility, inter-rater reliability and concurrent validity. One raised concerns regarding ceiling effects in an outpatient setting.[147] At the time of the review, only two articles had examined responsiveness of the UK FIM+FAM,[59, 147] and only one had examined the psychometric properties of the newer Extended Activities of Daily Living module - reporting inter-rater and test-retest reliability.[130] We found no previously published reports on the internal consistency or the factor structure of the UK FIM+FAM. We therefore undertook a further analysis of these aspects of its psychometric properties in a general neurorehabilitation sample.

## *PART 2: Scaling properties and dimensionality of the UK FIM+FAM in a mixed neurorehabilitation cohort*

### Participants and setting

Data were analysed from a single tertiary specialist in-patient rehabilitation service in London (catchment population in excess of five million) for patients with complex neurological disability. In this unit, the UK FIM+FAM had been routinely collected as part of routine clinical practice since 1999, although the EADL items were introduced gradually and only collected for all patients since August 2007.

The unit is the national training centre for the UK FIM+FAM, so that all staff received full training and regular updates on its application. UK FIM+FAM scores were routinely rated by the multidisciplinary treating team within 10 days of admission and during the last seven days before discharge. From a cohort of 764 consecutive patients admitted between January 1999 and December 2009, 459 had complete FIM+FAM data (including the EADL items) on admission and discharge. All 305 scores with missing EADL data were for admissions prior to August 2007. Between August 2007 and December 2009, data collection was complete for all admissions (n=188). Demographic characteristics of the cohort (n=459) were: Mean age 44.5 (SD 14) years; Males:Females 57:43%: Mean length of stay 101 (SD 61) days. Diagnosis: Acquired brain injury 384 (84%), strokes 256 (67%), traumatic 67 (17%), hypoxic/other 61 (16%), spinal cord injury 38 (8%), other neurological condition n=37 (8%), peripheral neurological conditions 33 (7%), progressive 4 (1%).

## Analysis

As this was the first examination of factor structure within the UK FIM+FAM, our two stage analysis included both exploratory and confirmatory components. FIM+FAM data are ordinal and often skewed, so we used a combination of parametric and non-parametric techniques. To do this we divided the sample at random (using SPSS v18 random sample selection) into two smaller samples of 225 (parametric, exploratory factor analysis) and 234 (non-parametric, Mokken scale analysis). For both samples we included each participant's admission and discharge FIM+FAM ratings, in order to maximise the range of ability sampled. This also doubled the sample size.

### Stage 1 – exploratory factor analysis

We first applied an exploratory factor analysis (EFA) to the pooled admission/discharge scores of the first sample (n=450). Even though they are based on parametric assumptions, principal components and factor analysis are widely used in this context, and have generally been considered appropriate for the initial stage of exploring and describing the relationships among a large set of variables, even where assumptions of normality may not strictly hold.[149] We followed the methods for EFA described earlier in this Chapter, using a principal components analysis with varimax rotation. This also allowed for direct comparison with the one previous factor analysis of the US FIM+FAM by Hawley et al 1999.[135, 150] On the basis of that previous factor analysis we rotated two components.[135] However, our own principal components analysis revealed four components with eigenvalues >1 (suggesting four substantial sources of variance), so we also examined a four-factor solution.

### Stage 2 – confirmatory Mokken analysis

In this particular evaluation we used a mixture ofparametric and non-parametric techniques to seeif the dimensions identified from the two methods were similar. We used Mokken analysis in our confirmatory analysis of the second pooled dataset (n=468). Mokken analysis is described earlier in this chapter.[64] In addition to examining the full 35 item scale, we also tested the subscales based on the two-factor and four-factor solutions provided by the EFA.

### Interpretation and Responsiveness

After completing EFA and CFA on the split dataset, internal consistency of the identified subscales was evaluated using Cronbach's alpha for the entire dataset (including both admission and discharge scores). In spite of the ordinal nature of FIM+FAM data, parametric and non-parametric evaluation of responsiveness in this large dataset gave very similar results. Here we report evaluation of responsiveness (change between admission and discharge) within the various subscales using paired t tests. Effect sizes are calculated using Cohen's d, taking account of the correlation between the means, and interpreted according to Cohen (0.2=Small, 0.5=Medium, 0.8=Large).[71] (A non-parametric analysis of responsiveness is available from the authors on request).

### Stage 1 - Exploratory Factor Analysis and Internal Consistency

The results of the principal components analysis with two factor Varimax rotations are presented in *Table 39.*

**Table 39: Principal Components Analysis with two- and four-factor varimax rotations of 30 FIM+FAM and 5 EADL Items (n=450)***

| ITEM | Median (IQR)* | Single Factor | Two Factors | | Four Factors | | | |
|---|---|---|---|---|---|---|---|---|
| | | 1st PC | Motor | Cognitive | Physical | Psycho-social | Commun-ication | EADL |
| Eating | 5 (5-7) | 0.79 | 0.52 | (0.61) | 0.57 | | | |
| Swallowing | 7 (6-7) | 0.65 | (0.34) | (0.58) | 0.51 | | (0.53) | |
| Grooming | 5 (4-7) | 0.89 | 0.63 | (0.63) | 0.62 | | | |
| Bathing | 4 (3-6) | 0.89 | 0.77 | | 0.73 | | | |
| Dressing: upper | 5 (3-7) | 0.88 | 0.68 | (0.55) | 0.66 | | | |
| Dressing: lower | 3 (2-6) | 0.87 | 0.84 | | 0.76 | | | |
| Toileting | 5 (2-7) | 0.84 | 0.85 | | 0.84 | | | |
| Bladder management | 6 (3-7) | 0.74 | 0.67 | | 0.74 | | | |
| Bowel management | 6 (3-7) | 0.72 | 0.64 | | 0.74 | | | |
| Transfers: bed/ chair | 5 (3-7) | 0.85 | 0.86 | | 0.86 | | | |
| Transfers: toilet | 5 (2-6) | 0.85 | 0.87 | | 0.87 | | | |
| Transfers: tub/shower | 4 (1-6) | 0.80 | 0.87 | | 0.81 | | | |
| Car transfer | 3 (1-5) | 0.79 | 0.85 | | 0.73 | | | |
| Locomotion | 5 (1-6) | 0.79 | 0.79 | | 0.75 | | | |
| Stairs | 1 (1-6) | 0.71 | 0.86 | | 0.78 | | | |
| Community mobility | 1 (1-3) | 0.74 | 0.75 | | | | 0.71 | |
| Comprehension | 6 (4-7) | 0.67 | 0.84 | | | | 0.73 | |
| Expression | 5 (3-7) | 0.68 | 0.84 | | | | 0.81 | |
| Reading | 5 (4-7) | 0.66 | 0.79 | | | | 0.70 | |
| Writing | 4 (2-6) | 0.70 | 0.77 | | | | 0.73 | |
| Speech intelligibility | 7 (4-7) | 0.59 | 0.65 | | | | 0.74 | |
| Social interaction | 6 (5-7) | 0.71 | 0.77 | | | 0.75 | | |
| Emotional status | 6 (3-7) | 0.63 | 0.61 | | | 0.68 | | |
| Adjustment | 5 (3-6) | 0.76 | 0.65 | | | 0.69 | | |
| Use of leisure time | 6 (3-6) | 0.81 | 0.67 | | | 0.51 | | |
| Problem solving | 5 (2-6) | 0.81 | 0.78 | | | 0.68 | | |
| Memory | 5 (3-7) | 0.75 | 0.77 | | | 0.73 | | |
| Orientation | 7 (4-7) | 0.75 | 0.80 | | | 0.72 | | |
| Concentration | 6 (4-7) | 0.75 | 0.76 | | | 0.74 | | |
| Safety awareness | 4 (2-6) | 0.82 | 0.67 | | | 0.60 | | |
| Meals | 2 (1-5) | 0.78 | 0.73 | | | | | 0.67 |
| Laundry | 1 (1-2) | 0.60 | 0.59 | | | | | 0.78 |
| Housework | 1 (1-2) | 0.61 | 0.66 | | | | | 0.77 |
| Shopping | 1 (1-3) | 0.68 | 0.68 | | | | | 0.79 |
| Financial management | 1 (1-3) | 0.57 | | 0.55 | | | | 0.58 |

**IQR = 25th -75th centiles: All items included the full a score range of 1-7; PC = Principal Component*

*EADL = Extended Activities of Daily Living*

*Note: All factor loadings rounded to 2 decimal points. Loadings < 0.50 removed for clarity.*
*The first two components extracted accounted for 66% of the total variance in responses.*
*All items loaded strongly on the first principal component (i.e. above 0.55).*

The first two components extracted accounted for 66% of the total variance in responses. All items loaded strongly on the first principal component (i.e. above 0.55). *Table 39* also shows a reasonably clear two-factor structure, with the 35 items falling into a *Motor* and a *Cognitive* factor. Within the EADL module, all items loaded on the Motor factor, with the exception of *Financial management* which loaded on the Cognitive factor. Similarly, the results of the four factor analysis also showed show four clear, interpretable factors corresponding to the following dimensions of independence: Physical independence, Psychosocial independence, Communication and Extended Activities of Everyday Living (EADL). In this solution, *Community Mobility* loaded onto the EADL factor.

A modest degree of overlap was seen for some items. The *Eating, Grooming* and *Dressing upper body* items loaded onto both Motor and Cognitive factors, and *Swallowing* loaded onto both the Physical and the Communication factors in the four-factor solution. For pragmatic reasons (and in line with the well-established Motor and Cognitive subscales of the FIM and FIM+FAM[135]) we elected to place all four of these items within the Motor and Physical parts of the scale. This led to the identification of:

1. Two principal 'domains'
   - *Motor*: 20 items (range score 20-140)
   - *Cognitive*: 15 items (range score 15-105) and

2. Four 'subscales'
   - *Physical*: 15 items (range score 15-105),
   - *Psychosocial*: 9 items (range score 9-63),
   - *Communication*: 6 items (range score 6-42) and
   - *EADL*: 5 items (range score 5-35).

These were then tested in the confirmatory analysis.

### Stage 2- Confirmatory Mokken Analysis

*Table 40* presents Loevinger's H coefficient for the overall scale and each individual item within each scale for the full 35-item scale and the subscales of the two- and four-factor solutions provided by the EFA.

The H coefficient of the full 35 item scale was 0.64 reflecting a strong scale.

- For the *Motor* and *Cognitive* scales the H coefficient was 0.82 and 0.65 respectively, once again reflecting strong scales. The H coefficient values for individual items were high across all three scales and always well above the accepted 0.30 cut-off.
- In the four-factor solution, once again H Coefficient values for each subscale were high, ranging from 0.67 – 0.82, indicative of strong scales. Individual item H coefficient values were also all high (i.e. >0.50) and all well above the accepted cut-off (i.e. H>0.30).

### Consistency

Consistency was tested for these across the whole dataset. The full scale reliability (internal consistency) was high with Cronbach's $\alpha$ = 0.98 for the full scale, and item-total correlations ranging from 0.56 – 0.88. Cronbach's alpha was 0.97 and 0.96 respectively for the *Motor* and *Cognitive* domains, and 0.97, 0.95, 0.92 and 0.90 respectively for the *Physical, Psychosocial, Communication* and *EADL* subscales.

### Responsiveness

Change in domain and subscale scores between admission and discharge is shown in *Table 41*. Significant changes were seen in all four subscales with 'large' effect sizes ranging from 0.86-1.29.

**Table 40: H Coefficient values from Mokken Analysis for 30 item FIM+FAM and 5 EADL items and for two- and four-factor based solutions (n=468)**

| ITEM | Median (IQR)* | Single Factor | Two Factors | | Four Factors | | | |
|---|---|---|---|---|---|---|---|---|
| | | Total scale | Motor | Cognitive | Physical | Psycho-social | Commun-ication | EADL |
| Eating | 5 (5-7) | 0.70 | 0.82 | | 0.82 | | | |
| Swallowing | 7 (6-7) | 0.73 | 0.78 | | 0.78 | | | |
| Grooming | 5 (4-7) | 0.69 | 0.80 | | 0.80 | | | |
| Bathing | 4 (3-6) | 0.71 | 0.85 | | 0.85 | | | |
| Dressing: upper | 5 (3-7) | 0.69 | 0.81 | | 0.81 | | | |
| Dressing: lower | 3 (2-6) | 0.69 | 0.85 | | 0.86 | | | |
| Toileting | 5 (2-7) | 0.68 | 0.86 | | 0.86 | | | |
| Bladder management | 6 (3-7) | 0.65 | 0.77 | | 0.77 | | | |
| Bowel management | 6 (3-7) | 0.63 | 0.77 | | 0.77 | | | |
| Transfers: bed/ chair | 5 (3-7) | 0.69 | 0.86 | | 0.86 | | | |
| Transfers: toilet | 5 (2-6) | 0.68 | 0.86 | | 0.86 | | | |
| Transfers: tub/shower | 4 (1-6) | 0.66 | 0.83 | | 0.83 | | | |
| Car transfer | 3 (1-5) | 0.66 | 0.82 | | 0.82 | | | |
| Locomotion | 5 (1-6) | 0.66 | 0.80 | | 0.80 | | | |
| Stairs | 1 (1-6) | 0.66 | 0.84 | | 0.84 | | | |
| Community mobility | 1 (1-3) | 0.70 | 0.79 | | | | | 0.68 |
| Comprehension | 6 (4-7) | 0.56 | | 0.68 | | | 0.73 | |
| Expression | 5 (3-7) | 0.56 | | 0.66 | | | 0.77 | |
| Reading | 5 (4-7) | 0.54 | | 0.64 | | | 0.72 | |
| Writing | 4 (2-6) | 0.54 | | 0.58 | | | 0.70 | |
| Speech intelligibility | 7 (4-7) | 0.52 | | 0.58 | | | 0.71 | |
| Social interaction | 6 (5-7) | 0.57 | | 0.65 | | 0.71 | | |
| Emotional status | 6 (3-7) | 0.48 | | 0.53 | | 0.58 | | |
| Adjustment | 5 (3-6) | 0.60 | | 0.68 | | 0.76 | | |
| Use of leisure time | 6 (3-6) | 0.66 | | 0.68 | | 0.72 | | |
| Problem solving | 5 (2-6) | 0.64 | | 0.72 | | 0.77 | | |
| Memory | 5 (3-7) | 0.56 | | 0.67 | | 0.75 | | |
| Orientation | 7 (4-7) | 0.63 | | 0.69 | | 0.76 | | |
| Concentration | 6 (4-7) | 0.59 | | 0.65 | | 0.73 | | |
| Safety awareness | 4 (2-6) | 0.63 | | 0.71 | | 0.75 | | |
| Meals | 2 (1-5) | 0.68 | 0.77 | | | | | 0.76 |
| Laundry | 1 (1-2) | 0.69 | 0.78 | | | | | 0.67 |
| Housework | 1 (1-2) | 0.69 | 0.80 | | | | | 0.68 |
| Shopping | 1 (1-3) | 0.68 | 0.75 | | | | | 0.72 |
| Financial management | 1 (1-3) | 0.58 | | 0.71 | | | | 0.52 |
| **Scale H coefficient** | | **0.64** | **0.82** | **0.65** | **0.82** | **0.72** | **0.72** | **0.67** |

**IQR = 25th-5th centiles. All items included the full score range of 1-7; EADL=Extended Activities of Daily Living*

**Table 41: Scales and domain scores on admission and discharge, and change scores (n=459)**

| | Admission | | Discharge | | Mean diff | 95% CI | Paired T tests | | | Effect size |
|---|---|---|---|---|---|---|---|---|---|---|
| | Mean (SD) | Range | Mean (SD) | Range | | | t | df | p | |
| **Subscales** | | | | | | | | | | |
| **Physical** | 55.2 (26.0) | 15-105 | 77.2 (26.7) | 15-105 | 22.1 | 20.5, 23.6 | 27.6 | 457 | <0.001 | 1.29 |
| **Psychosocial** | 40.0 (15.9) | 9-63 | 46.9 (14.5) | 9-63 | 7.0 | 6.2, 7.7 | 18.1 | 457 | <0.001 | 0.86 |
| **Communication** | 22.9 (9.2) | 5-35 | 26.6 (8.1) | 5-35 | 3.7 | 3.3, 4.1 | 18.2 | 458 | <0.001 | 0.87 |
| **EADL** | 8.8 (5.1) | 6-41 | 16.7 (9.1) | 6-42 | 7.9 | 7.2, 8.5 | 22.7 | 458 | <0.001 | 1.21 |
| **Domains** | | | | | | | | | | |
| **Motor** | 63.9(28.0) | 20-133 | 85.8 (30.7) | 19-132 | 22.0 | 20.3, 23.6 | 26.2 | 457 | <0.001 | 1.24 |
| **Cognitive** | 64.7 (24.4) | 15-105 | 76.3 (22.6) | 15-105 | 11.7 | 10.7, 12.7 | 22.0 | 457 | <0.001 | 1.05 |

*EADL = Extended Activities of Daily Living; CI=Confidence Interval*

**In summary,** the systematic review of existing literature supported the psychometric robustness of the UK FIM+FAM, although as with the US version, ceiling effects can be a problem with outpatient samples. The analysis of new data from a large, mixed neurorehabilitation cohort demonstrated that the UK FIM+FAM has a highly acceptable level of internal consistency or reliability. Moreover the internal consistency was high not simply for the full 35 item scale, but also for the two *Motor* and *Cognitive* domains and the four subscales (*physical, psychosocial, communication, EADL*) similarly identified by factor analysis. This suggests that the FIM+FAM is a particularly useful measure as it can be used to derive a reliable, single score of overall independence, but also yields specific information on four separate dimensions of independence. Mokken analysis confirmed that each of these constituted a reliable, unidimensional ordinal scale appropriate for rank ordering persons.

In this analysis we used a combination of parametric and non-parametric approaches for exploratory and confirmatory analysis, which to our knowledge is novel. If a plurality of approaches yields the same conclusions, it increases the likelihood that the findings are robust. This approach may also have application in other areas of rehabilitation measurement, where clinical data are typically ordinal and often are not normally distributed.

### Limitations of the study

- The analyses were completed on patients from a single tertiary rehabilitation service. so the results require replication in an independent population.
- Using both admission and discharge scores ensured representation across the full score range for all items in both samples but may have inflated the degree of internal consistency giving an enhanced impression of homogeneity.
- We were not able to test the more recently added '*Work*' item in the EADL, as there was insufficient representation in this dataset. This will require evaluation in future analyses.

Notwithstanding the acknowledged limitations, our findings demonstrated that the UK FIM+FAM is a reliable and responsive measure of independent functioning suitable for use in mixed inpatient neurorehabilitation settings. Whilst it may be acceptable to sum item scores into a single total figure, the instrument also provides meaningful scores on two and four sub-dimensions.

This study also provided the first examination of the relationship between the 30 FIM+FAM items and the 5 additional EADL items. The results suggested that these five items provide useful information on these more extended activities of everyday living and that they combined well with the existing 30 FIM+FAM items, but can also be used as a standalone module.

A summary of the psychometric properties of the US and UK FIM+FAM according to the medical Outcomes Trust framework is given in Appendix 4.9.

## *Analysis and presentation of FIM+FAM data*

As noted above, both the FIM and the UK FIM+FAM generate ordinal level data and there is continuing debate about how such data should be handled for analysis. Because the FIM+FAM is the principal outcome measure within the UKROC database, it is pertinent to examine this question in more detail.

## *Analysis of item level data*

At item level, the various scale cut-off points often have useful clinical meaning for rehabilitation professionals, which often need to be retained for clinical decision-making.

For example, *Figure 23* illustrates a FAM-Splat for an individual patient following stroke. The 'FAM splat' represents a radar chart of FIM+FAM scores:

- The 30 items are arranged as the 'spokes' of a wheel, each being rated on a scale of 1 (totally dependent) to 7 (fully independent).
- The blue-shaded area illustrates the change from admission to discharge.
- The dotted lines represent goals scores that were not fully achieved during the rehabilitation programme.

Item scores of 6 or above typically indicate that the individual is able to function independently (with or without an aid or device), whilst scores of 5 or below indicate increasing requirements for help from another person.

In this example it can be seen at a glance that:

- By discharge from rehabilitation, the patient has made substantial change from near total dependence for self-care activities and mobility on discharge to nearly complete independence,
- He has also recovered full independence for communication and cognition.
- He still requires assistance from a carer for bathing, getting around in the community and pursuing leisure activities - as well as some continued support for managing his emotions and adjustment to limitations. However, in all other areas, he can manage independently providing he has the appropriate aids and adaptations within his home environment.

In addition to providing an 'at-a-glance' view of the functional ability of individual patients, the FAM splat can be used to illustrate the profile of different groups of patients by plotting the median scores for the group on a 'composite FAM splat'.

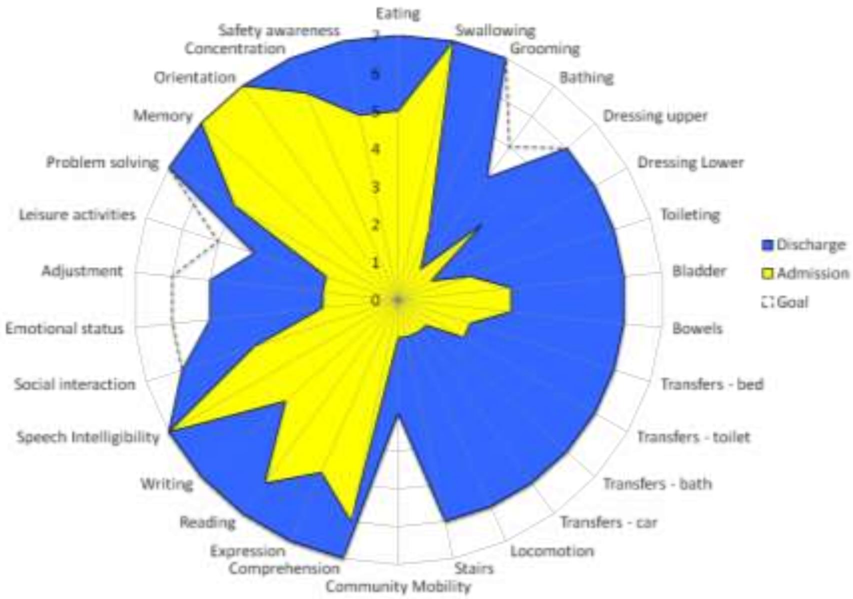



**Figure 23:** **FAM splat of a patient recovering from Stroke**

Drawn from an analysis of the UKROC dataset in January 2013, *Figure 24* shows composite FAM splats representing the median scores for populations of 6085 patients with acquired brain injury (ABI) and 1509 patients with progressive neurological conditions (such as Multiple sclerosis, Parkinson's disease etc).

As would be expected the pattern of disability is quite different, the brain injured patients having more severe cognitive and communicative problems than patients with progressive neurological conditions. Nevertheless, the value of the additional FIM+FAM items is identifiable, even in the non brain-injured population as only 6 items of function are unaffected.

**Acquired Brain Injury** **Progressive condition**

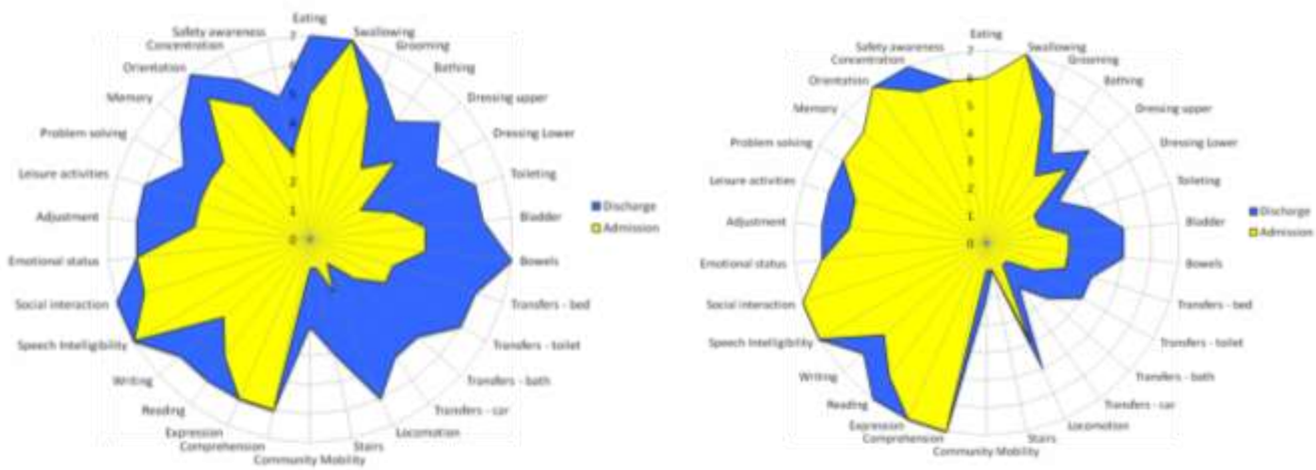


**Figure 24:** **Composite FAM splats for populations of acquired brain injury and progressive conditions**

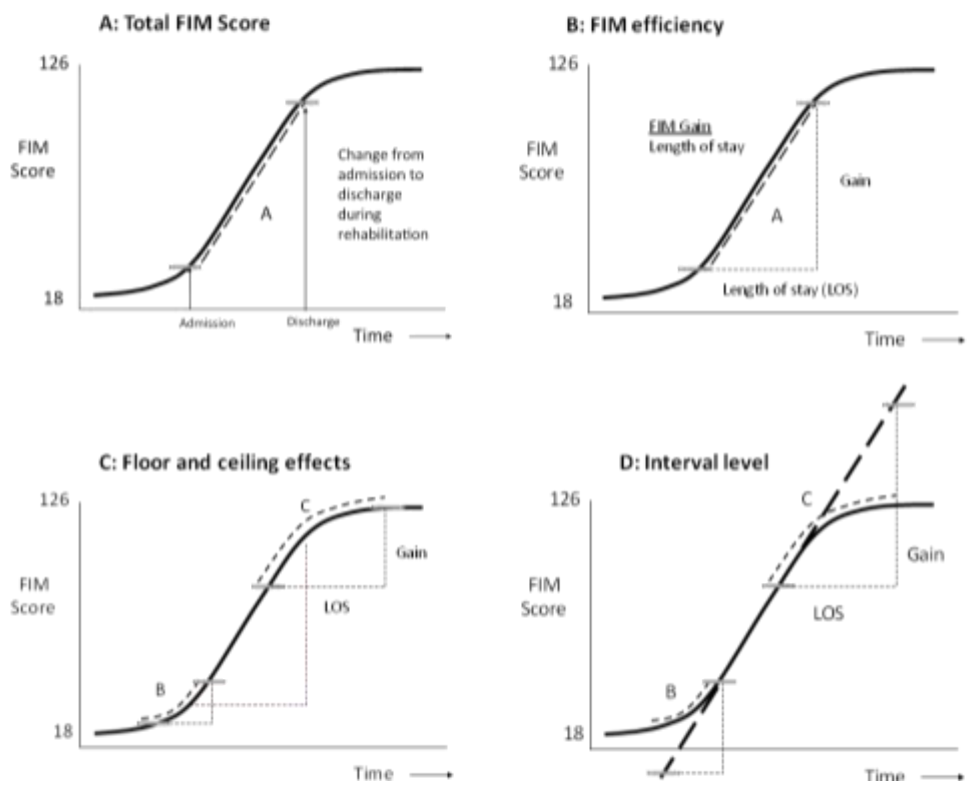
From these FAM splats one can see that, as a group, the population of patients with progressive conditions requires more assistance for self care and is largely wheelchair bound, although modest gains were made in rehabilitation – especially with respect to bladder and bowel management. However, communication, cognitive and psychosocial function are relatively intact, compared with the brain injury patients who still require most assistance from others in these areas. Thus, it is helpful to preserve the integrity of these ordinal level ratings for their clinical meaning, whether describing an individual patient or a population.

## *Analysis of summary data*

These ordinal scores, however, do not produce interval level data. Such data may be analysed using non-parametric statistics (treating data as 'ranks' rather than actual numbers) but should not be strictly treated as if they have true numerical value – they should not be added together or divided by something else.

Within clinical practice, many ordinal scales are in fact found to be additive and summed data are frequently used to provide an overall reflection of the parameter of interest. When the numbers are large and normally distributed, parametric statistical analysis will often produce very similar results to non-parametric analysis.

A common feature of many ordinal scales, however, is that they typically demonstrate a curvi-linear progression over time, with floor and ceiling effects at either end of the scale as illustrated in *Figure 25.* Case-mix systems in the US and Australia use FIM efficiency (Gain in total FIM score ÷ length of stay) as a surrogate measure of cost efficiency. However, as is clear from this illustration, if this mathematical manipulation of ordinal data is valid at all, it can only reasonably be used for data in the 'straight line' part of the scale. For this reason many authors have advocated transformation of ordinal scale to interval level data to provide more accurate and meaningful measurement. A particular advantage of transformation is that it can help to reduce floor and ceiling effects and improve sensitivity at either end of the scale.




***Legend to Figure 25:***

*Figure 26 illustrates the change in total Functional Independence Measure (FIM) score during rehabilitation demonstrating the typical curvi-linear progression over time*

- *3a, Patient A is admitted to rehabilitation at the beginning of the straight part of the curve, and discharged before the FIM score starts to plateau.*
- *3b demonstrates the calculation of 'FIM efficiency' (FIM gain/length of stay),*
- *3c illustrates the floor and ceiling effects of the FIM respectively for Patient B, admitted with profound disability and patient C who is mobile but in post-traumatic amnesia.*
- *3d illustrates how conversion to interval level data can potentially provide a better reflection of change, especially at the upper and lower limits of the scale.*

**Figure 25: Change in Functional Independence Measure (FIM) total score during rehabilitation**

## *Transformation to ordinal level data – Rasch Analysis*

Developed by the Danish mathematician, Georg Rasch, in the 1960s[151] Rasch analysis uses item response theory to provide a probabilistic, logistic model which postulates that achieving a particular test score is influenced by just two variables – the ability of the person and the difficulty of the test item. The model produces logit values (locations) describing both 'item difficulty' and 'person ability', in order to transform ordinal scores to a genuine interval scale.

An extensive literature has explored Rasch analysis of FIM data and produced a variety of interval-level transformation models for the FIM scale.[152] To date, however, the FIM+FAM has received little exposure to Rasch modelling. Two previous studies have explored the benefits of Rasch transformation of the original US version in patients following stroke[133] and traumatic brain injury[135] but as yet there have been no published Rasch analyses of the UKFIM+FAM in any population.

Work undertaken as part of this programme has explored Rasch analysis on a population of 320 patients following stroke. Before broadening our analysis to neurological conditions in general, we wished to determine with the UK FIM+FAM fits the Rasch model for this relatively discrete population. Acknowledging the fairly clear clinical differences in function that may be expected between left and right hemisphere strokes, we were also interested to know whether Rasch-transformed interval level scales would show discriminant levels of function between left and right strokes, as had previously been demonstrated using the ordinal scores.[153]

**Methods:** As developments in Rasch methodology are evolving rapidly, we undertook two independent analyses. Both analyses were conducted on exactly the same dataset extracted from the UKROC database.

**Sampling:** We extracted the cohort of all 1318 stroke patients consecutively admitted to the 58 Level 1/2 specialist rehabilitation centres in England between January 1, 2010 and May 30, 2013, for whom a complete UK FIM+FAM score was available at both admission and at discharge from the unit.

- FIM+FAM scores are expected to be lower on admission and higher at discharge from rehabilitation. To ensure that the data represented the full range of the scale, we pooled admission and discharge scores from the complete sample of N=1318, into one dataset.
- In order not to violate the Rasch assumption of local independence between observations (i.e. to prevent the same patient contributing two entries in the data) we included only one time point, i.e. admission or discharge, for each patient.
- Taking into account the largest sub-division of the UK FIM+FAM identified from previous factor analyses (i.e. the 16 motor items) we used a randomly selected sample of 320 cases (representing 20 cases per item for this domain) to fulfil the sampling criteria.[133] The Rasch analysis was then performed using this sample.

**Demographics:** Within our random sample of 320 cases, the male: female ratio was 59:41%; the mean age was 58.7 years (SD=15.27 range 16 to 89); and the mean length of stay was 78.9 (52.6) days. To confirm that this group was representative of the cohort from which it was drawn, we compared the socio-demographic and clinical characteristics of our Rasch study sample with the full cohort (see *Table 42*). No significant differences were seen.

For this analysis, our data included only the 30-item FIM+FAM scale (physical, psychosocial and communication subscales) and not the EADL module.

**Table 42: The UKROC stroke population and the Rasch random sample characteristics**

| Demographics | UKROC study sample N=1318 | | Random sample (Rasch analysis)* N=320 | |
|---|---|---|---|---|
| Length of stay (days) | Mean<br>77.7 | SD<br>(57.3) | Mean<br>78.9 | SD<br>(52.6) |
| Gender, male | n<br>752 | (%)<br>(57.1) | n<br>189 | (%)<br>(59.1) |
| **Age, years** | **n** | **%** | **n** | **%** |
| < 44 | 220 | (16.7) | 50 | (15.6) |
| 45-54 | 293 | (22.2) | 74 | (23.1) |
| 55-64 | 298 | (22.6) | 66 | (20.6) |
| 65-74 | 250 | (19.0) | 54 | (16.9) |
| 74+ | 231 | (17.5) | 68 | (21.3) |
| Unknown | 26 | (2.0) | 8 | (2.5) |
| **Ethnicity** | **n** | **%** | **n** | **%** |
| White | 951 | (72.2) | 227 | (70.9) |
| Asian/Asian British | 98 | (7.4) | 21 | (6.6) |
| Black/Black British | 110 | (8.3) | 29 | (9.1) |
| Other | 41 | (3.1) | 10 | (3.2) |
| Unknown | 118 | (8.9) | 33 | (10.3) |
| **Diagnosis localisation** | **n** | **%** | **n** | **%** |
| Right hemisphere | 638 | (48.4) | 159 | (49.7) |
| Left Hemisphere | 680 | (51.6) | 161 | (50.3) |
| **Diagnosis subcategory** | **n** | **%** | **n** | **%** |
| Haemorrhagic | 386 | (29.3) | 93 | (29.1) |
| Infarct | 707 | (53.6) | 174 | (54.4) |
| Sub-Arachnoid | 136 | (10.3) | 32 | (10.0) |
| Other | 89 | (6.8) | 21 | (6.6) |

** Random sample extracted from the dataset (n=1318) derived across admission and discharge values so that each patient is only in the dataset once but both time points are equally represented*

A preliminary factor analysis was conducted on both the total extracted sample (n=1318) and the subsample for Rasch analysis. Exploratory principal components analysis with varimax rotation suggested a 3-factor solution (Motor, and Cognitive subscales – the latter subdividing into Communication and Psychosocial function). This structure is in line with our previous studies,[132, 153] and explained 72% of the total variance in the responses within this sample.

# Analysis 1:

The initial exploration was conducted in the UK led by Dr Roxana Alexandrescu in 2013/14.
We followed the analysis protocols that were considered best practice in the literature at that time and as taught and advocated by the Rasch group of the University of Leeds.
http://medhealth.leeds.ac.uk/info/732/psychometric_laboratory/1489/rasch_courses

We first examined three subscales based upon the above factor analysis (Motor, Communication, Psychosocial) and then examined four based on clinical judgement (subdividing the Motor subscale into two components - Self-care and Mobility). The unaltered scale showed a significant misfit to the Rasch Model.

Within successive steps followed by consultation of fit statistics:

- The items with disordered thresholds were rescored from seven to a lower number of categories (19 items),
- Any items showing residual correlations > 0.3 were combined into testlets: 10 items were combined into five testlets: a) Eating and Swallowing; b) Dressing upper and lower body; c) Bladder and Bowel management; d) Transfer bed and toilet; e) Reading and Writing.
- We also split those items exhibiting statistically significant differential item functioning (DIF) by:
  - stroke hemispheric location (three items, Grooming, Expression, Concentration),
  - gender (Toileting) and
  - age (Social interaction, testlet Reading and Writing).

The modifications are summarised in *Table 44,* and full details of the rescoring schedule are given in *Table A, Appendix 4.14*. Draft ordinal/interval score conversion tables were drawn up, splitting the FIM+FAM into four factors and separating for left/right localization, age and gender as appropriate (see *Table B, Appendix 4.14)*.

**Table 43: The UK item level Rasch model fit of the scale**

| Item | (scale) | Disordered threshold | DIF | Testlet | Location (order by location) | Fit residual | P value |
|---|---|---|---|---|---|---|---|
| **Motor self-care domain** | | | | | | | |
| Swallowing | (FAM) | yes | no | yes [c] | -1.332 (1) | 4.079 | 0.036 [b] |
| Eating | (FIM) | yes | no | yes [c] | -1.332 (1) | 4.079 | 0.036 [b] |
| Grooming | (FIM) | no | R/L | no | -0.721/-0.259 [a] (2/3) | -0.516/-0.603 | 0.31/0.29 |
| Bowel management | (FIM) | yes | no | yes [c] | -0.172 (4) | 3.284 | 0.55 |
| Bladder management | (FIM) | yes | no | yes [c] | -0.172 (4) | 3.284 | 0.55 |
| Dressing lower | (FIM) | no | no | yes [c] | 0.324 (5) | -2.921 | 0.36 |
| Dressing upper | (FIM) | no | no | yes [c] | 0.324 (5) | -2.921 | 0.36 |
| Bathing | (FIM) | no | no | no | 0.539 (6) | -1.827 | 0.06 |
| Toileting | (FIM) | yes | gender F/M | no | 0.706/0.916 [a] (7/8) | -0.876/-3.321 | 0.42/0.006[b] |
| **Motor mobility domain** | | | | | | | |
| Transfers bed | (FIM) | no | no | yes [c] | -2.064 (1) | -4.312 | 0.79 |
| Transfers toilet | (FIM) | no | no | yes [c] | -2.064 (1) | -4.312 | 0.79 |
| Transfers bath | (FIM) | yes | no | no | -0.471 (2) | -0.572 | 0.005 |
| Locomotion | (FIM) | yes | no | no | 0.208 (3) | 0.510 | 0.93 |
| Transfers car | (FAM) | yes | no | no | 0.264 (4) | 0.006 | 0.04 |
| Community mobility | (FAM) | yes | no | no | 0.614 (5) | 2.170 | 0.007 |
| Stairs | (FIM) | yes | no | no | 1.448 (6) | -1.962 | 0.06 |
| **Communication domain** | | | | | | | |
| Comprehension | (FIM) | no | no | no | -0.843 (1) | -1.062 | 0.26 |
| Speech | (FAM) | no | no | no | -0.769 (2) | 2.869 | 0.47 |
| Expression | (FIM) | yes | R/L | no | -.137/0.693[a] (3/5) | -0.427/-0.726 | 0.04/0.007 |
| Reading | (FAM) | yes | age | yes [c] | 0.209/0.846[a] (4/6) | -0.726/-0.195 | 0.58/0.98 |
| Writing | (FAM) | yes | age | yes [c] | 0.209/0.846[a] (4/6) | -0.726/-0.195 | 0.58/0.98 |
| **Psychosocial domain** | | | | | | | |
| Social interaction | (FIM) | yes | age | no | -0.501/-1.005[a] (4/1) | 0.609/0.020 | 0.73/0.97 |
| Concentration | (FAM) | no | R/L | no | -0.082/-0.755[a] (6/2) | 0.180/-0.206 | 0.15/0.24 |
| Orientation | (FAM) | yes | no | no | -0.678 (3) | -1.927 | 0.11 |
| Emotional status | (FAM) | yes | no | no | -0.193 (5) | 2.134 | 0.18 |
| Memory | (FIM) | no | no | no | -0.072 (7) | 1.223 | 0.66 |
| Adjustment | (FAM) | no | no | no | 0.205 (8) | 0.709 | 0.22 |
| Problem solving | (FIM) | no | no | no | 0.649 (9) | -2.575 | 0.02[b] |
| Safety awareness | (FAM) | yes | no | no | 1.147 (10) | -0.671 | 0.27 |
| Leisure activities | (FAM) | yes | no | no | 1.285 (11) | -0.441 | 0.30 |

*The UK FIMFAM: UK Functional Independence Measure and Functional Assessment Measure. [a] DIF by localisation Right/ Left stroke for Grooming, Expression, Concentration; gender female/male for Toileting; and age 0-54y/55y+ for testlet Reading and Writing; Social interaction. In bold fit residual >2.5 or <-2.5 (if >2.5 indicates the item measures a different concept than the rest of the domain; if<-2.5 indicates local dependency; redundancy); [b] non-significant, probability below values for Bonferroni adjustment (i.e., 0.006 motor self care; 0.0017 motor mobility; 0.0017 communication; 0.0045 psychosocial); [c] maximum score is the sum of the maximum scores of the individual items involved. Rescore structure in Rasch is by default 0123456. Location (order by location): location of the item on the Rasch logit ruler from the easiest item (negative value) to the most difficult item (positive value) and the corresponding ranking value based on the position on the scale.*

*Figure 26* shows scatterplots of the transformed versus the rescored ordinal scores for the four subscales. Despite the calculated differential item functioning, the differences between left and right strokes and gender and age were very minimal for the Motor self care and psychosocial scales, suggesting that for practical purposes these two subscales could each be condensed to a single transformation column.

The communication subscale, however, did show a difference between left and right strokes, and also between older and younger patients. These differential factors were therefore retained in the simplified transformation table that was drawn up for clinical purposes. Highlighted in yellow, see *Table B, Appendix 4.14*.

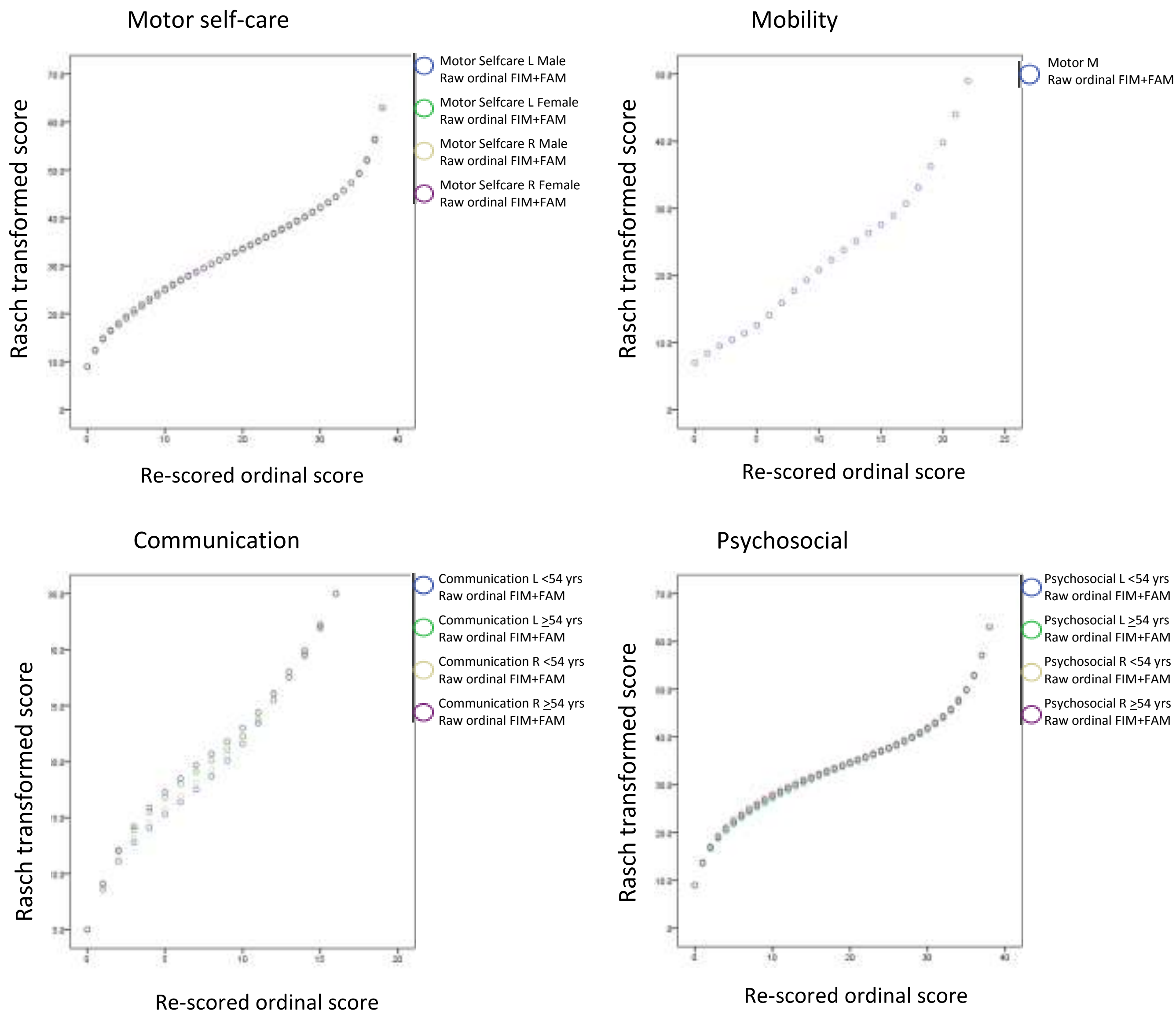



**Figure 26: Scatterplots of the transformed versus re-scored ordinal scores for the four subscales – Analysis 1.**

Table 44 shows a comparison of data using the full and simplified tables to transform the re-scored FIM+FAM subscale scores. The final transformed scores (shown in red) are seen to be very similar.

The differential item functioning for communication between left and right strokes is not unexpected as a significantly greater proportion of patients with left hemisphere strokes would be expected to be dysphasic and so have poorer function in this domain. Therefore, the results did resonate to some extent with clinical expectation.

Even with these considerable modifications, however, the fit was marginal and some evidence of multi-dimensionality remained. Although T-tests for uni-dimensionality showed a good fit for the Mobility and Communication testlets at 1.79% and 0.74% respectively (desirable <5%), they were marginal for Self-care (6.25%) and Psychosocial (5.88%). The overall fit statistics for the 4-subscale model are shown in Table 45.

**Table 44: Comparison of transformation scores using the full and simplified versions of the transformation table for Analysis 1 (see Table B, Appendix 4.10)**

| | Left Hemisphere | | | | | | Right Hemisphere | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Sub-scales** | **Admission** | | | **Discharge** | | | **Admission** | | | **Discharge** | | |
| | Total Re-scored | Transf'd Full | Transf'd Simple | Total Re-scored | Transf'd Full | Transf'd Simple | Total Re-scored | Transf'd Full | Transf'd Simple | Total Re-scored | Transf'd Full | Transf'd Simple |
| Self-care | *14* | 28.8 | 28.7 | *28* | 40.2 | 40.3 | *14* | 28.7 | 28.7 | *24* | 36.7 | 36.8 |
| Mobility | *4* | 11.4 | 11.4 | *15* | 27.6 | 27.6 | *4* | 11.4 | 11.4 | *11* | 22.3 | 22.3 |
| Communication | *7* | 19.1 | 19.1 | *10* | 22.3 | 22.3 | *11* | 24.4 | 24.4 | *12* | 26.1 | 26.1 |
| Psychosocial | *17* | 32.4 | 32.8 | *26* | 38.1 | 38.4 | *17* | 32.8 | 32.8 | *28* | 40 | 40 |
| Total Motor | | 40.2 | 40.1 | | 67.8 | 67.9 | | 40.1 | 40.1 | | 59.0 | 59.1 |
| Total Cognitive | | 51.5 | 51.9 | | 60.4 | 60.7 | | 57.2 | 57.2 | | 66.1 | 66.1 |
| Total Motor % | | 35.9% | 35.8% | | 60.5% | 60.6% | | 35.8% | 35.8% | | 52.7% | 52.8% |
| Total Cognitive % | | 52.6% | 53.0% | | 61.6% | 61.9% | | 58.4% | 58.4% | | 67.4% | 67.4% |

***"Transf'd full"*** *uses the transformed scores from the full version of the transformation table, separating the Motor Self-care scale for localisation (i.e. Left/Right) and gender, and separating the Psychosocial scale for localization and age.*
***"Transf'd simple"*** *uses the simplified version highlighted in yellow, applying just a single transformation scale for each of 'Motor self-care' and 'Psychosocial' function.*

**Table 45: Overall fit of the UK FIM FAM to the Rasch model - summary statistics for Analysis 1**

| UK FIM FAM domain Rasch model | Item Location [a] | Item Fit residual [a] | Person Location [a] | Person Fit residual [a] | Item –Trait Interaction | | PSI | Unidimensionality t-test (%) |
|---|---|---|---|---|---|---|---|---|
| | | | | | χ square/DF | p value | | |
| **Motor** | | | | | | | | |
| Motor self-care | 0.000 (0.768) | -0.338 (2.694) | 0.467 (2.180) | -0.351 (0.974) | 56.725/32 | 0.005 | 0.91 | 6.25 |
| Motor mobility | 0.000 (1.189) | -0.693 (2.230) | -1.183 (3.034) | -0.369 (0.806) | 50.212/24 | 0.001 | 0.93 | 1.79 |
| **Cognition** | | | | | | | | |
| Communication | 0.000 (0.715) | -0.052 (1.462) | 0.244 (1.921) | -0.331 (0.914) | 36.314/24 | 0.05 | 0.83 | 0.74 |
| Psychosocial | 0.000 (0.759) | -0.086 (1.338) | 0.412 (1.772) | -0.269 (0.999) | 58.818/44 | 0.07 | 0.91 | 5.88 |

*UK FIM FAM: UK Functional Independence Measure and Functional Assessment Measure*
*[a] Mean (Standard Deviation) ; DF: degrees of freedom; PSI: Person Separation Index;*
*The values for perfect fit to the model: fit residual 0.0 (SD=1); p value > 0.01; PSI>0.8; t test <5%.*
*Location of the item is set to 0.00; location of the person should be closer to the item location for a well-targeted scale.*
*The overall mean values for perfect fit for both item and person fit residual are 0.0 (SD=1).*

## *Analysis 2*

The second analysis was completed by the team in New Zealand in 2015, led by Professors Richard Siegert and Paula Kersten, with assistance from Oleg Medvedev. This analysis was conducted on the advice of Prof Alan Tennant (grant collaborator and internal expert on Rasch Analysis) in line with the latest Rasch methods for 2015.

This approach treated the 30 items as a single scale.

- To retain sensitivity, disordered thresholds were only re-scored if the 95% confidence intervals did not overlap (see *Figure 27*). Just 8 disordered thresholds were re-scored (for details of the re-scoring schedule, see *Table C, Appendix 4.14*).
- We then created three testlets representing the Motor, Communication and Psychosocial domains. This approach resulted in excellent fit to the Rasch model (unidimensionality 1.88%). However, some local dependency was seen between the Communication and Psychosocial testlets.
- After combining these, unidimensionality was still within acceptable limits (4.88%), however, the item-trait interaction was now significant (p<0.01) with evidence of differential item functioning for localisation of stroke (see *Table 46*).
- After splitting both testlets into left and right-hemisphere strokes, an excellent fit was obtained with high reliability (Person separation Index 0.88 and unidimensionality 4.88% with no local dependency.

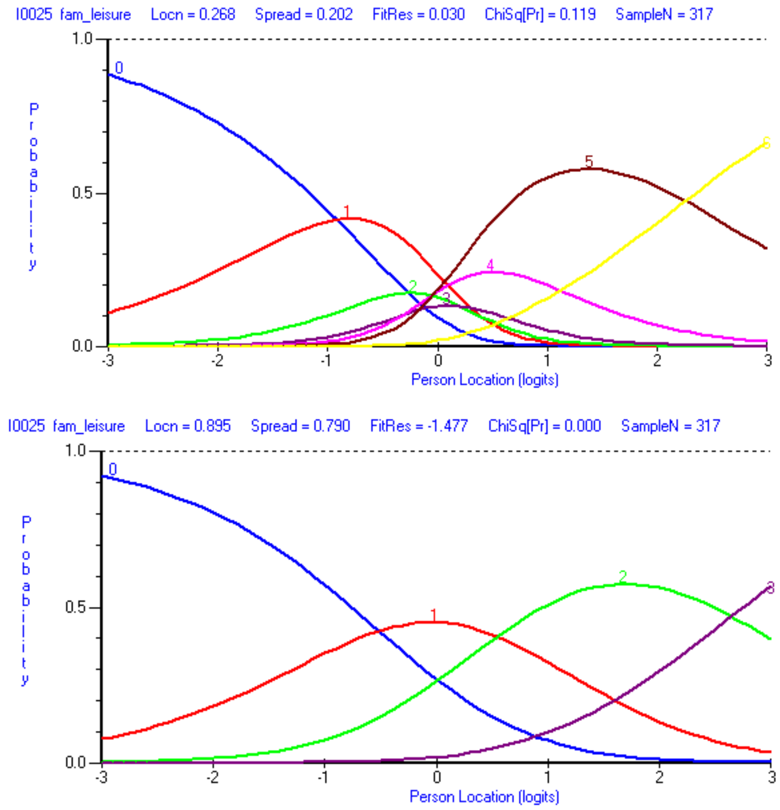



**Figure 27: Item category probability curves illustrating disordered thresholds for the FIM+FAM Leisure activities item (before re-scoring (top panel) and orderly thresholds after re-scoring (bottom panel).**

**Table 46: Overall fit of the UK FIM FAM to the Rasch model - summary statistics for Analysis 2**

| UK FIM FAM domain Rasch model | Item Location [a] | Item Fit residual [a] | Person Location [a] | Person Fit residual [a] | Item –Trait Interaction | | PSI | Unidimensionality t-test (%) |
|---|---|---|---|---|---|---|---|---|
| | | | | | χ square/DF | p value | | |
| **Analysis 1 (3 Factors)** | 0.000 (0.041) | 0.171 (1.620) | 0.058 (0.279) | -0.327 (0.875) | 18.565/12 | 0.09 | 0.80 | 1.88 |
| **Analysis 2 (2 Factors)** | 0.000 (0.035) | 0.058 (1.168) | 0.036 (0.206) | -0.337 (0.703) | 20.008/8 | 0.01 | 0.72 | 4.38 |
| **Final Analysis (2 Factors)** | 0.000 (0.291) | 0.147 (0.673) | 0.0487 (0.323) | -0.328 (0.718) | 19.017/16 | 0.27 | 0.88 | 4.38 |

*UK FIM FAM: UK Functional Independence Measure and Functional Assessment Measure*
*[a] Mean (Standard Deviation) ; DF: degrees of freedom; PSI: Person Separation Index;*

Separate ordinal/interval score conversion tables were therefore created for left and right hemisphere strokes in just two domains (Motor and Cognitive) – see Table D Appendix 4.14. As the original scale totals are 112 and 98 respectively, we rounded both scales to a range of 0-100 in this analysis.

*Figure 28* shows a scattergram of Rasch-transformed full scale scores versus raw ordinal FIM+FAM scores (after re-scoring) in patients with left and right stroke.

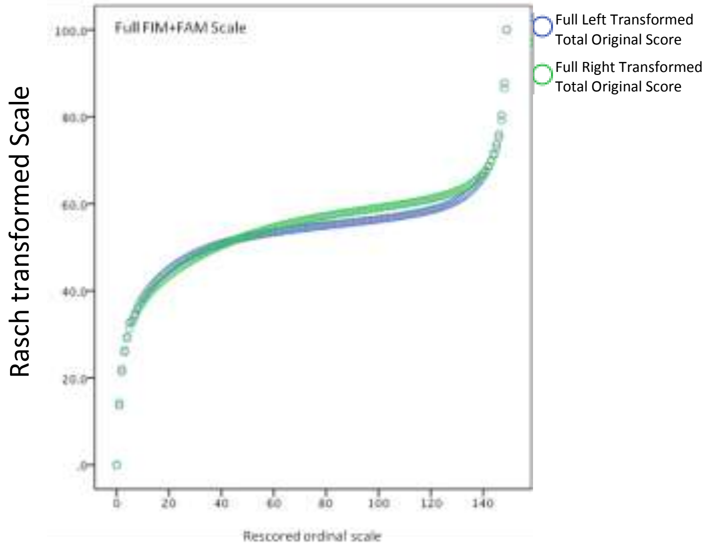



**Figure 28: Scattergram of Rasch-transformed full scale scores versus raw (re-scored) FIM+FAM scores**

For the total scale, there was relatively little separation of left and right strokes. However, the motor and cognitive scales showed greater differential item functioning for left and right strokes, as illustrated in *Figure 29.*

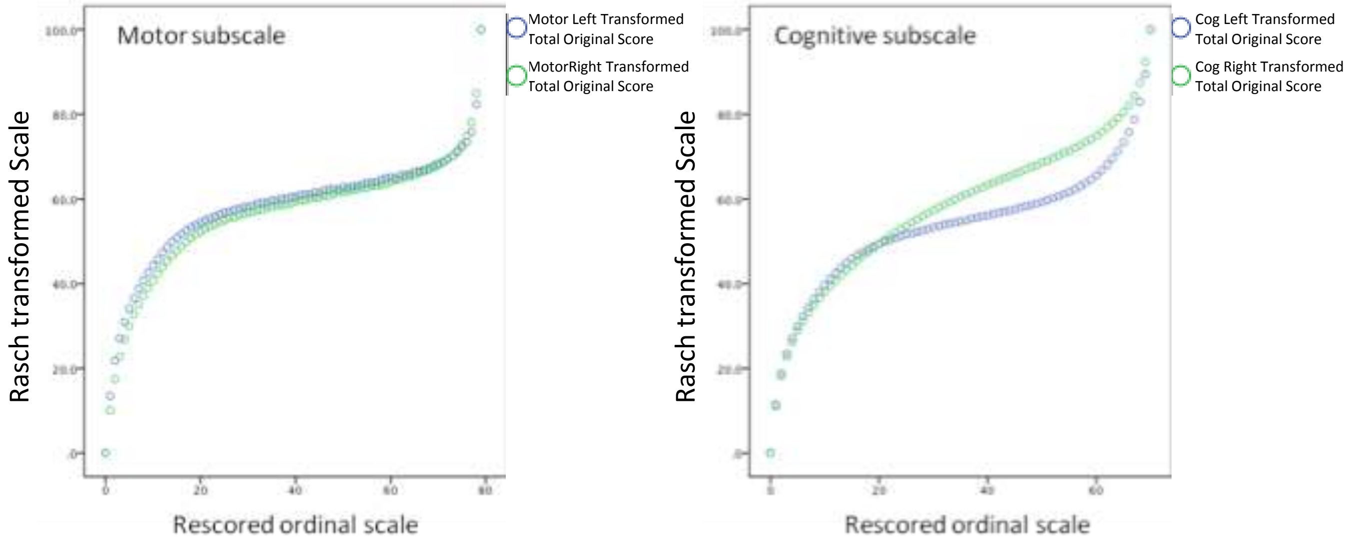



**Figure 29: Scattergram of Rasch-transformed scores versus raw (re-scored) FIM+FAM scores for the motor and cognitive scales**

Although the curve is smooth and the summary statistics suggest that the UK FIM+FAM fits the Rasch model very well, the central part of the curve is notably flat. This suggests that the interval-level transformed scale may be relatively insensitive in the middle range of the scale. This may potentially limit the usefulness of the interval-level UK FIM+FAM as an outcome measure for clinical practice.

Of the two analyses, the slope in the middle range of the scale appeared to be steeper for the first analysis than the second. However, this could simply be a scaling effect, as the score ranges (i.e. 9-63, 7-49, 5-35 and 9-63) were narrower for the 4-subscale solution (Analysis 1) than for the 2-subscale solution (0-100 each) in Analysis 2.

We were therefore uncertain which (if either) of these models might have more practical utility. Before setting up a complex computing algorithm to derive the transformed scores, we therefore wished to examine the potential impact of these different approaches.

Re-scoring and Rasch transformation were conducted in a test analysis to examine the conversion of FIM+FAM data from a live sample and to compare the two interval models with the original raw FIM+FAM scores.

## Test analysis

For this analysis we used the stroke sub-sample from the April 2010-2015 extract of the UKROC dataset. Data for all episodes of rehabilitation following stroke in a Level 1 or 2 specialist rehabilitation unit were extracted if they met the following criteria:

- Length of stay >7 and ≤400 days (i.e. admission for assessment only and long stay were excluded).
- A valid FIM+FAM score was recorded both at admission and discharge.

This yielded a dataset of n=3623 episodes: Males:females 51:49%, mean age 58 (SD 15.7) years and mean length of stay 85 (SD 59 days).

Item-level median FIM+FAM scores were calculated. *Figure 30* shows the FAM splat of the median scores on admission and discharge, whilst *Figure 31* illustrates the difference between median FIM+FAM scores for right and left strokes on admission and discharge.

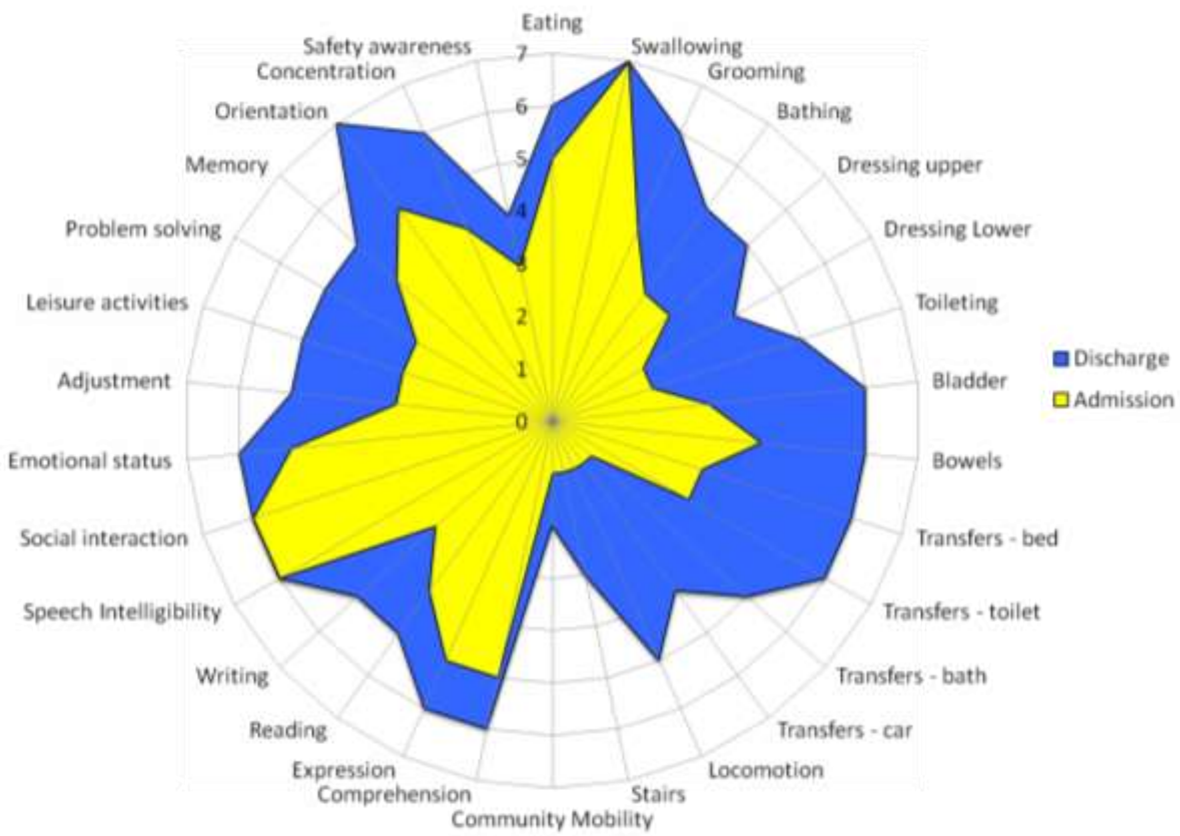
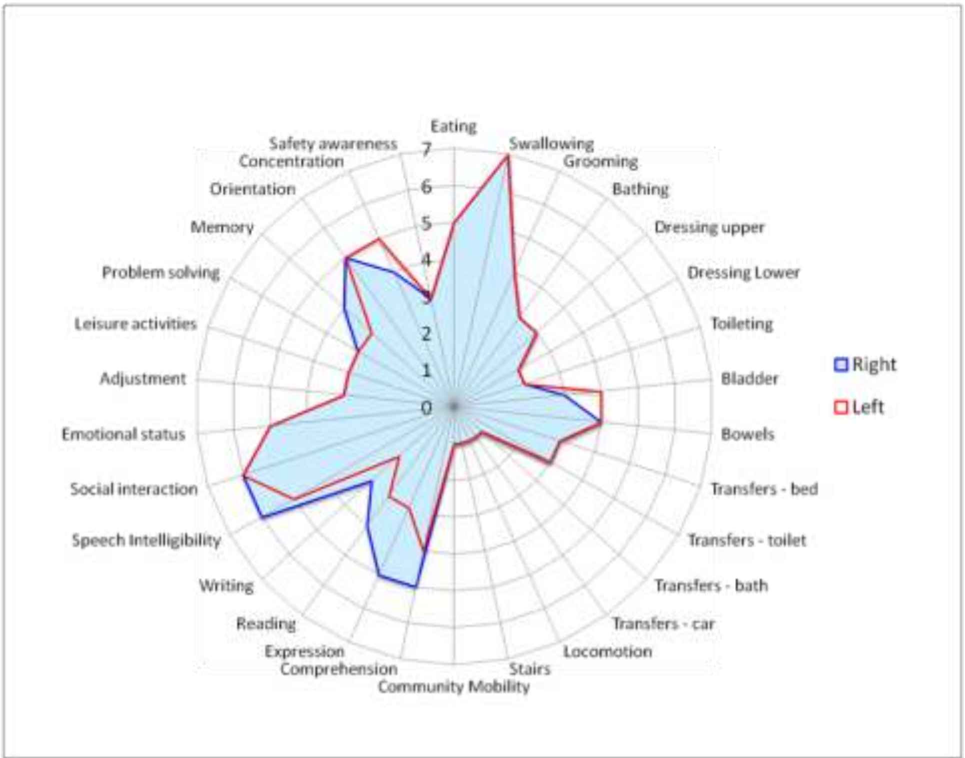



**Figure 30:** **FAM splat of median scores on admission and discharge for the stroke population N=3623**



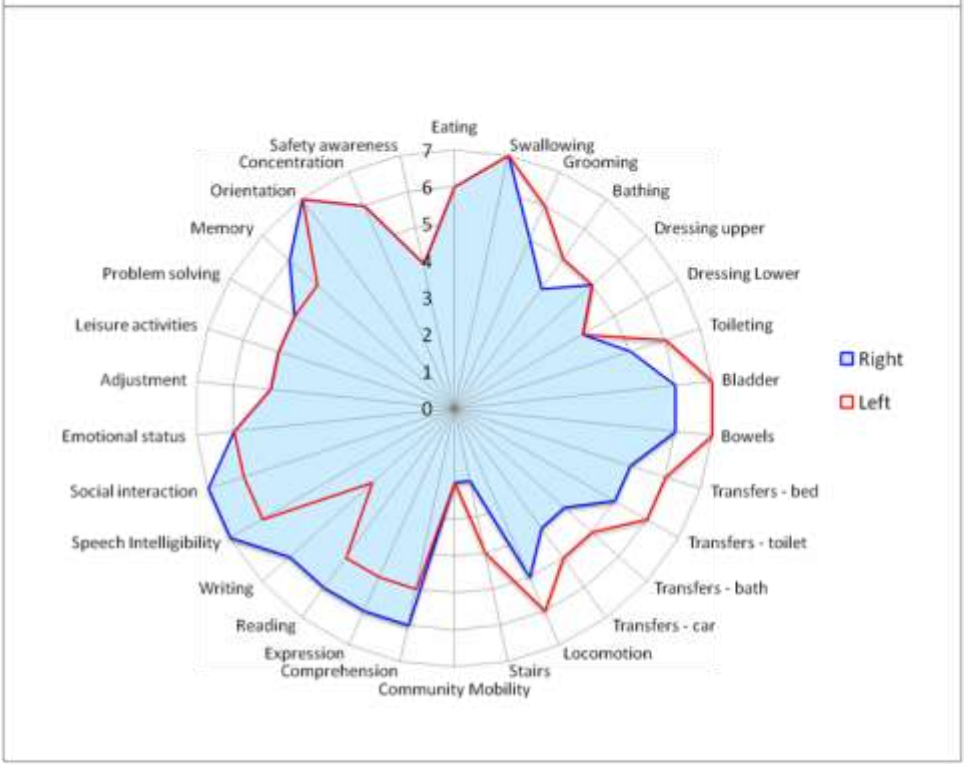




**Figure 31: FAM splats of the median scores for left (n=1370) and right (n=1277) stroke populations on admission (top) and discharge (bottom)**

*Table 47* shows the results of independent sample parametric (T tests) and non-parametric (Mann Whitney) statistical analysis of the difference between left and right hemisphere stroke for the various domains and subscales of the FIM+FAM on both admission and discharge.

**Table 47: T tests and Mann Whitney U tests for subscale sub-scale differences on admission and discharge for left (n=1370) and right (n=1277) hemisphere strokes**

| | T tests | | | | | | Mann Whitney | |
|---|---|---|---|---|---|---|---|---|
| | t | Mean Difference | 95% CIs Upper | Lower | P value 2 tailed | | z | P value 2 tailed |
| **Admission** | | | | | | | | |
| Self-care | 3.7 | 1.6 | 0.7 | 2.4 | <0.001 | | 3.52 | <0.001 |
| Sphincters | 2.0 | 0.4 | 0.0 | 0.7 | 0.051 | | 1.797 | 0.072 |
| Transfers | 5.8 | 1.7 | 1.1 | 2.3 | <0.001 | | 5.088 | <0.001 |
| Locomotion | 6.2 | 1.1 | 0.8 | 1.5 | <0.001 | | 4.361 | <0.001 |
| Communication | -20.7 | -6.7 | -7.4 | -6.1 | <0.001 | | -19.336 | <0.001 |
| Psychosocial | -4.0 | -1.0 | -1.4 | -0.5 | <0.001 | | -3.836 | <0.001 |
| Cognition | -4.3 | -1.4 | -2.1 | -0.8 | <0.001 | | -4.197 | <0.001 |
| **Subscales** | | | | | | | | |
| FAM Motor | 4.8 | 4.8 | 2.8 | 6.7 | <0.001 | | 3.939 | <0.001 |
| FAM Cognitive | -11.5 | -9.1 | -10.7 | -7.5 | <0.001 | | -11.227 | <0.001 |
| **FAM Total** | **-2.8** | **-4.3** | **-7.4** | **-1.3** | **0.005** | | **-2.887** | **0.004** |
| | | | | | | | | |
| **Discharge** | | | | | | | | |
| Self-care | 3.4 | 1.5 | 0.6 | 2.3 | 0.001 | | 4.132 | <0.001 |
| Sphincters | 2.7 | 0.4 | 0.1 | 0.8 | 0.007 | | 3.237 | <0.001 |
| Transfers | 5.5 | 1.8 | 1.1 | 2.4 | <0.001 | | 5.617 | <0.001 |
| Locomotion | 5.4 | 1.2 | 0.8 | 1.7 | <0.001 | | 5.376 | <0.001 |
| Communication | -19.8 | -5.8 | -6.4 | -5.3 | <0.001 | | -19.403 | <0.001 |
| Psychosocial | -2.9 | -0.6 | -1.1 | -0.2 | 0.003 | | -2.802 | 0.005 |
| Cognition | -2.7 | -0.8 | -1.4 | -0.2 | 0.007 | | -2.326 | 0.02 |
| **Subscales** | | | | | | | | |
| FAM Motor | 4.6 | 4.9 | 2.8 | 7.0 | <0.001 | | 4.995 | <0.001 |
| FAM Cognitive | -9.8 | -7.3 | -8.8 | -5.8 | <0.001 | | -10.181 | <0.001 |
| **FAM Total** | **-1.5** | **-2.4** | **-5.7** | **0.8** | **0.144** | | **-1.284** | **0.199** |

*Red = Left hemisphere lower; Blue = Right hemisphere lower; Black = non significant*

Although the scores had improved significantly between these two time points, significant differences were observed between left and right strokes, with similar patterns being evident at both admission and discharge.

- The FIM+FAM cognitive scales was lower for the left hemisphere stroke population - the difference being principally reflected in the communication domain.
- The right hemisphere stroke population had significantly lower scores in the motor scale although the differences were smaller and fairly equally spread across the self-care, transfers and locomotion domains, and most probably suggesting a degree of dyspraxia.

The parametric and non-parametric tests yielded almost identical results. The rescoring and calculation of Rasch-transformed scores were applied to the median item scores.

In Table 48 and Figure 32 we compare the original raw scores, re-scored raw scores and Rasch-transformed scores. To make them comparable the mean scores are expressed as a % of the total score. For both analyses (-but particularly Analysis 2) the Rasch-transformed change scores can be seen to represent a significantly smaller proportion of the total score than the raw scores (with or without re-scoring).

**Table 48: Comparison of results for raw, re-scored and Rasch transformed FIM+FAM scores calculated for median item scores for left (n=1370) and right (n=1277) stokes**

| Series | | Left | | | Right | | |
|---|---|---|---|---|---|---|---|
| | | Admission | Discharge | Change | Admission | Discharge | Change |
| **Raw FIM +FAM scores** | | | | | | | |
| | **Motor** | 40.2% | 77.7% | **37.5%** | 40.2% | 67.0% | **26.8%** |
| | **Cognitive** | 54.1% | 74.5% | **20.4%** | 66.3% | 83.7% | **17.3%** |
| **Re-scored raw scores** | | | | | | | |
| **Analysis 1** | **Motor** | 30.0% | 71.7% | **41.7%** | 30.0% | 58.3% | **28.3%** |
| | **Cognitive** | 44.4% | 66.7% | **22.2%** | 51.9% | 74.1% | **22.2%** |
| **Analysis 2** | **Motor** | 32.9% | 74.7% | **41.8%** | 32.9% | 62.0% | **29.1%** |
| | **Cognitive** | 47.1% | 68.6% | **21.4%** | 61.4% | 80.0% | **18.6%** |
| **Rasch transformed scores** | | | | | | | |
| **Analysis 1** | **Motor** | 35.8% | 60.6% | **24.8%** | 35.8% | 52.8% | **17.0%** |
| | **Cognitive** | 53.0% | 61.9% | **8.9%** | 58.4% | 67.4 % | **9.0%** |
| **Analysis 2** | **Motor** | 57.0% | 64.9% | **8.3%** | 67.6% | 74.4% | **6.8%** |
| | **Cognitive** | 50.3% | 54.0% | **4.3%** | 67.3% | 73.9% | **6.6%** |

Figure 32 illustrates the patterns for the raw, rescored and Rasch-transformed scores for the two analyses.

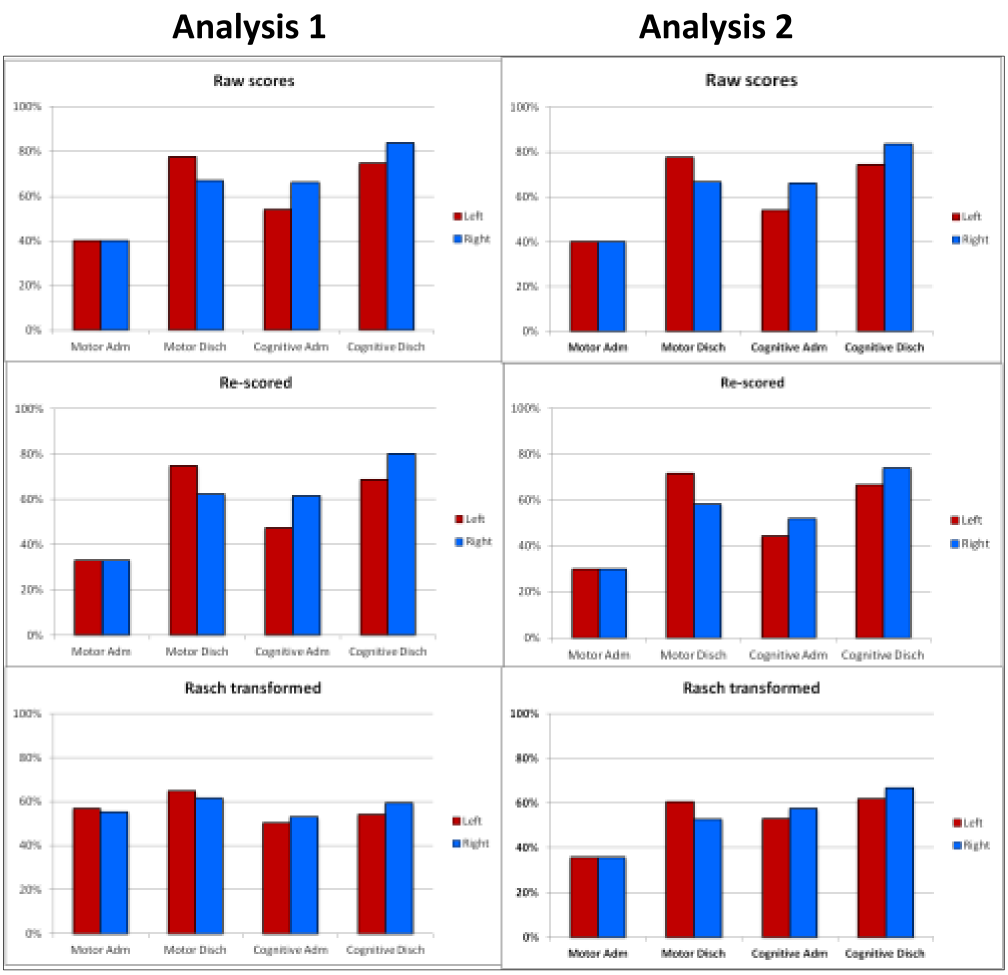



**Figure 32: Patterns for the raw, rescored and Rasch-transformed scores for the two analyses**

# Discussion

Of the two Rasch analyses, Analysis 2 provided a better fit to the Rasch Model in this population, the scale dividing into two principal subscales – a 16-item Motor scale and 14-item Cognitive scale, with differential item functioning for left and right strokes.  However, in order to produce this good fit, 10 outlying right hemisphere stroke patients had to be excluded.

Analysis 1 provided a statistically less good fit, and a somewhat more cumbersome transformation table, although we found that this could be simplified for clinical purposes without significant loss of sensitivity.

That said, however, while the Rasch-transformed scores may increase the sensitivity at the extreme ends of the scale, they tended to give a flatter profile in the middle range – particularly Analysis 2. Thus, in percentage terms, only approximately a quarter of the change between admission and discharge that is apparent on the raw and re-scored ordinal scales was seen for the transformed scores in Analysis 2. Analysis 1 faired slightly better for the motor component of the scale, but still fell short in the cognitive elements.

In addition, whilst the mean and median raw scores demonstrate differences between left and right strokes that are expected clinically, the transformed scores were less sensitive to these differences, especially in Analysis 2.

And although both analyses demonstrated the UK FIM+FAM to fit the Rasch model within acceptable limits, the two methods conducted on an identical dataset provided quite strikingly different results, which tends to reduce overall confidence in the approach.

As noted above, Rasch methodology is still evolving. The 2011 review by Lundgren and Tennant of over 50 articles that applied Rasch analysis to the FIM[152] demonstrated the wide variation in results from applying different approaches.  Developments even within the timescale of this programme have led to a significant change in approach which, as demonstrated here, can impact substantially on the end result.

So whilst we accept the principle that, in an ideal world, it is desirable to use interval level data for analysis, the quest for methodological purity should not be at the expense of clinical utility.

In this particular context, the ordinal scores represent a language that has clinical meaning for rehabilitation professionals, and its practical interpretation can be a useful aid to clinical decision-making. For example, the FAM splats provide an 'at-a-glance' understanding of which daily tasks a patient requires assistance from another person, and which they can manage for themselves given the right equipment.  A comparative analysis showed that parametric and non-parametric statistical tests of differences between domain and subscale scores yielded almost identical results.

We conclude that, in its current state of development, Rasch-transformed FIM+FAM scores do not provide sufficient benefit over the raw ordinal scores to allow us to recommend it. Interestingly, a similar conclusion was reached by Hawley et al 1999[135] for the US FIM+FAM.

# Outcomes – The Barthel Index (BI)

Originally described by Mahoney and Barthel in 1963[154] the Barthel Index is a simple but widely used measure of independence in basic activities of daily living. Several different versions now exist, with rating scales ranging from 0-20 to 0-100. The UKROC database incorporates the Barthel Index (BI) based on the manual published by Collin and Wade in 1988,[155] with a score range of 0-20.

The minimum dataset required of Level 2 services includes the RCS (version 8) as a measure of complexity and the BI as a measure of outcome. The BI may be rated directly, but can also be derived from either the NPDS[156] or the FIM[86] and the algorithms for these conversions are incorporated in the UKROC software. Thus the BI can potentially form a basic common language outcome measure for any service that records the BI, FIM±FAM or the NPDS.

In clinical rehabilitation settings, the NPDS is normally rated by the nursing staff, whilst the UK FIM+FAM should be rated by the multidisciplinary team (MDT). In theory, the MDT should include a member of the nursing team when rating the FIM+FAM, but in practice this is not always possible. Therefore, the NPDS tends to reflect a nursing perspective of the patient's performance whilst the FIM±FAM represents the therapists' perspective.

The common language of the BI provides an opportunity to compare directly these two perspectives as captured in the course of routine clinical practice, and it is also important to understand the extent to which these two potential sources of BI data may differ, and why.

In the following cohort analysis we compared BI ratings as derived from:

- NPDS scores rated by the nursing staff (NPDS-BI)
- FIM scores – rated mainly by the therapy team (FIM-BI)

We expected to find a reasonably close relationship, but not an exact one as patients frequently perform differently in day-to-day activities with nursing staff compared with their performance during therapy sessions.

## *Methods*

From the UKROC Dataset – April 2010 to April 2014, we extracted all cases that had both valid NPDS and FIM scores on admission. A total of 4887 cases were extracted representing data from 52 specialist rehabilitation services across England.

We examined the relationship between the NPDS-BI and the FIM-BI scores, both on admission and discharge. As the dataset was large and the Barthel scores near-normally distributed, parametric statistics were used alongside non-parametric statistics.

- Correlations between the total and individual items scores were tested with Pearson correlation tests.
- Significant differences in item and total scores were sought using paired t tests.
- Agreement between total scores was tested by intra class correlation coefficients (two-way random effects model) and the Bland Altman 95% limits of agreement (mean difference ± 1.96 x SD).
- Item by item agreement was tested using unweighted Cohen's kappa coefficients with 95% confidence intervals, interpreted according to Landis and Koch.[157]
- Items with kappa coefficients <0.3 were further explored through cross-tabulation.

## *Results*

### Demographics

The mean age of the sample was 54.4 years (SD16.6), male:female ratio 58:42%. The mean length of stay was 82 (SD 69) days. In total, 3467 (71%) of patients had acquired brain injury, 565 (12%) had spinal cord injury, 590 (12%) had a progressive condition and 265 (5%) had peripheral neurological conditions.

### Comparison of ratings on admission

Descriptives for the NPDS-BI and FIM-BI on admission are shown in *Table 49*.

**Table 49: Descriptive statistics for NPDS-BI and FIM-BI on admission (N=4887)**

| | NPCNA conversion (NPDS-BI) | | | | FIM conversion (FIM-BI) | | | |
|---|---|---|---|---|---|---|---|---|
| **Item** | Mean | SD | Median (IQR) | Range | Mean | SD | Median (IQR) | Range |
| **Bowels** | 1.25 | 0.88 | 2 (0-2) | 0-2 | 1.2 | 0.94 | 2 (0-2) | 0-2 |
| **Bladder** | 0.89 | 0.95 | 0 (0-2) | 0-2 | 1.34 | 0.91 | 2 (0-2) | 0-2 |
| **Grooming** | 0.24 | 0.43 | 0 (0-0) | 0-1 | 0.23 | 0.42 | 0 (0-0) | 0-1 |
| **Toilet use** | 0.54 | 0.79 | 0 (0-1) | 0-2 | 0.52 | 0.74 | 0 (0-1) | 0-2 |
| **Feeding** | 1.06 | 0.83 | 1 (0-2) | 0-2 | 1.2 | 0.82 | 1 (0-2) | 0-2 |
| **Transfers** | 1.25 | 1.17 | 1 (0-2) | 0-3 | 1.23 | 1.21 | 1 (0-2) | 0-3 |
| **Mobility** | 0.84 | 1.11 | 0 (0-2) | 0-3 | 0.67 | 1.10 | 0 (0-1) | 0-3 |
| **Dressing** | 0.32 | 0.64 | 0 (0-0) | 0-2 | 0.84 | 0.61 | 1 (0-1) | 0-2 |
| **Stairs** | 0.27 | 0.60 | 0 (0-0) | 0-2 | 0.23 | 0.59 | 0 (0-0) | 0-2 |
| **Bathing** | 0.06 | 0.23 | 0 (0-0) | 0-1 | 0.07 | 0.26 | 0 (0-0) | 0-1 |
| **Total BI** | **6.71** | **5.68** | **5 (2-11)** | **0-20** | **7.52** | **5.58** | **7 (3-11)** | **0-20** |

**Correlation and agreement** between the two sources for total and individual item scores are shown in *Table 50*.

**Table 50: Spearman rho correlations between total and individual item scores**

| **Item** | **rho** | **Significance** | **Kappa** | **95% CI** | **Interpretation** |
|---|---|---|---|---|---|
| **Bowels** | 0.61 | p<0.001 | 0.47 | 0.45, 0.49 | Moderate |
| **Bladder** | 0.35 | p<0.001 | 0.29 | 0.27, 0.30 | Fair |
| **Grooming** | 0.41 | p<0.001 | 0.41 | 0.38, 0.44 | Moderate |
| **Toilet use** | 0.69 | p<0.001 | 0.52 | 0.50, 0.54 | Moderate |
| **Feeding** | 0.69 | p<0.001 | 0.50 | 0.48, 0.52 | Moderate |
| **Transfers** | 0.78 | p<0.001 | 0.54 | 0.52, 0.55 | Moderate |
| **Mobility** | 0.73 | p<0.001 | 0.56 | 0.54, 0.58 | Moderate |
| **Dressing** | 0.51 | p<0.001 | 0.18 | 0.17, 0.19 | Poor |
| **Stairs** | 0.50 | p<0.001 | 0.39 | 0.36, 0.42 | Fair |
| **Bathing** | 0.43 | p<0.001 | 0.48 | 0.43, 0.52 | Moderate |
| | | | **ICC** | | |
| **Total BI** | **0.86** | **p<0.001** | **0.87** | **0.86, 0.88** | |

The correlation between total scores was rho 0.86, intra class coefficient (ICC) 0.87 (95% CI 0.86, 0.88). The 95% limits of agreement ranged from -4.8 to 6.5.

Individual item correlations ranged from 0.20 to 0.60 (all significant at p<0.001). Kappa coefficients ranged from 0.18 to 0.56, showing 'moderate' agreement for 7/10 items. However, they were below 0.3 for 2 items – 'Bladder (urinary continence) and 'Dressing'.

**Significant differences** between the two sets of ratings are shown in *Table 51*.

**Table 51: Paired T-tests and Wilcoxon signed rank for the differences between NPDS-BI and FIM-BI scores**

| Item | Mean Difference | 95% CIs | | Paired T tests | | Wilcoxon | |
|---|---|---|---|---|---|---|---|
| | | Lower | Upper | t | p | z | p |
| Bowels | -0.05 | -0.07 | -0.03 | -4.43 | **<0.001** | **-4.0** | **<0.001** |
| Bladder | 0.45 | 0.42 | 0.48 | 29.79 | **<0.001** | **-27.7** | **<0.001** |
| Grooming | -0.01 | -0.02 | 0.00 | -1.72 | 0.085 | -1.7 | 0.085 |
| Toilet use | -0.03 | -0.04 | -0.01 | -3.10 | 0.002 | -3.2 | 0.002 |
| Feeding | 0.14 | 0.12 | 0.16 | 15.10 | **<0.001** | **-14.7** | **<0.001** |
| Transfers | -0.02 | -0.04 | 0.01 | -1.40 | 0.162 | -0.51 | 0.611 |
| Mobility | -0.17 | -0.19 | -0.15 | -16.60 | **<0.001** | **-15.9** | **<0.001** |
| Dressing | 0.52 | 0.51 | 0.54 | 60.56 | **<0.001** | **-11.7** | **<0.001** |
| Stairs | -0.05 | -0.06 | -0.03 | -5.56 | **<0.001** | **-24.4** | **<0.001** |
| Bathing | -0.02 | -0.02 | -0.01 | -4.78 | **<0.001** | **-4.8** | **<0.001** |
| **Total BI** | **0.81** | **0.73** | **0.89** | **19.54** | **<0.001** | **-20.2** | **<0.001** |

Overall, the NPDS-Derived BI was significantly lower than the FIM-Derived BI, but the direction of difference varied from item to item. Significant differences were seen between 7-10 items – three of these being lower for the NPDS-BI and four lower for the FIM-BI.

In general the differences were small, reaching statistical significance only because of the large sample size. However, the Bladder (urinary incontinence) and Dressing items showed larger differences, both of which were lower for the NPDS-derived rating, and probably at a level to be clinically significant.

### Further Investigation through cross tabulation for Bladder and Dressing

Cross tabulation of scores for these items on admission are shown in *Table 52* below.

**Table 52: Cross tabulation of NPDS-BI and FIM-BI ratings for Urinary continence and Dressing on admission**

| **Bladder** | **NPDS-BI** | | | |
|---|---|---|---|---|
| **FIM-BI** | 0 | 1 | 2 | **Total** |
| 0 | 1111 | 144 | 199 | 1454 |
| 1 | 152 | 72 | 110 | 334 |
| 2 | 1235 | 215 | 1648 | 3098 |
| **Total** | 2498 | 431 | 1957 | 4886 |

| **Dressing** | **NPDS-BI** | | | |
|---|---|---|---|---|
| **FIM-BI** | 0 | 1 | 2 | **Total** |
| 0 | 1323 | 8 | 8 | 1339 |
| 1 | 2358 | 461 | 150 | 2969 |
| 2 | 99 | 174 | 306 | 579 |
| **Total** | 3780 | 643 | 464 | 4887 |

In 1235 cases (25%) urinary continence was rated as 2 (fully continent) according to the therapists' FIM rating but 0 (incontinent) according to the nurses – meaning that the patient was either catheterised or had at least one accident per day. In 2358 (48%) of cases, Dressing was rated as 1 (needing some help) by the therapists but 0 (dependent) by the nurses.

## Comparison of ratings on discharge

By discharge the direction of difference was systematically in one direction. The NPDS-BI was now significantly lower than the FIM-BI on all but one item (mobility), so that the mean difference in total scores had increased to 1.66 - the 95% limits of agreement limits of agreement ranging from -1.4 to 4.7 (see *Table 53)*

However, on item-by item analysis, the differences were actually quite small, so that the agreement between the two methods for individual items had improved. The ICC for the total scores was 0.91 (95% CI 0.91, 0.92). – see *Table 54*

**Table 53: Descriptive statistics for NPDS-BI and FIM-BI on discharge**

| N=4887 | NPCNA conversion | | | | FIM conversion | | | |
|---|---|---|---|---|---|---|---|---|
| **Item** | **Mean** | **SD** | **Median (IQR)** | **Range** | **Mean** | **SD** | **Median (IQR)** | **Range** |
| Bowels | 1.47 | 0.81 | 2 (1-2) | 0-2 | 1.56 | 0.79 | 2 (1-2) | 0-2 |
| Bladder | 1.21 | 0.93 | 2 (0-2) | 0-2 | 1.59 | 0.78 | 2 (2-2) | 0-2 |
| Grooming | 0.49 | 0.50 | 0 (0-1) | 0-1 | 0.5 | 0.50 | 1 (0-1) | 0-1 |
| Toilet use | 0.97 | 0.90 | 1 (0-2) | 0-2 | 1.1 | 0.88 | 1 (0-2) | 0-2 |
| Feeding | 1.35 | 0.78 | 2 (1-2) | 0-2 | 1.52 | 0.72 | 2 (1-2) | 0-2 |
| Transfers | 1.91 | 1.17 | 1 (1-3) | 0-3 | 2.03 | 1.19 | 3 (1-3) | 0-3 |
| Mobility | 1.49 | 1.24 | 1 (0-3) | 0-3 | 1.47 | 1.32 | 1 (0-3) | 0-3 |
| Dressing | 0.72 | 0.84 | 0 (0-2) | 0-2 | 1.22 | 0.67 | 1 (1-2) | 0-2 |
| Stairs | 0.51 | 0.75 | 0 (0-1) | 0-2 | 0.74 | 0.87 | 0 (0-2) | 0-2 |
| Bathing | 0.17 | 0.38 | 0 (0-0) | 0-1 | 0.28 | 0.45 | 0 (0-1) | 0-1 |
| **Total BI** | **10.28** | **6.50** | **11 (4-16)** | **0-20** | **11.99** | **6.56** | **13 (7-18)** | **0-20** |

**Table 54: Spearman correlation and agreement (Kappa) between NPDS-BI and FIM-BI scores on discharge**

| Item | rho | Significance | Kappa | 95% CI | Interpretation |
|---|---|---|---|---|---|
| Bowels | 0.66 | p<0.001 | 0.52 | 0.50, 0.54 | Moderate |
| Bladder | 0.49 | p<0.001 | 0.38 | 0.36, 0.41 | Fair |
| Grooming | 0.59 | p<0.001 | 0.59 | 0.56, 0.61 | Moderate |
| Toilet use | 0.77 | p<0.001 | 0.59 | 0.57, 0.61 | Moderate |
| Feeding | 0.73 | p<0.001 | 0.53 | 0.51, 0.55 | Moderate |
| Transfers | 0.86 | p<0.001 | 0.63 | 0.61, 0.64 | Substantial |
| Mobility | 0.83 | p<0.001 | 0.62 | 0.60, 0.63 | Substantial |
| Dressing | 0.68 | p<0.001 | 0.27 | 0.25, 0.29 | Fair |
| Stairs | 0.61 | p<0.001 | 0.45 | 0.43, 0.47 | Moderate |
| Bathing | 0.55 | p<0.001 | 0.53 | 0.50, 0.56 | Moderate |
| | | | **ICC** | | |
| **Total BI** | **0.66** | **p<0.001** | **0.91** | **0.91, 0.92** | |

As on admission, the largest differences were seen for urinary continence and dressing, with Kappas <0.4 for these two items (see *Table 55*).

**Table 55: Paired T-tests for the differences between NPDS-BI and FIM-BI scores on discharge**

| Item | Mean Difference | 95% CIs | | Paired T tests | | Wilcoxon | |
|---|---|---|---|---|---|---|---|
| | | Lower | Upper | t | p | z | p |
| Bowels | 0.09 | 0.07 | 0.11 | 9.12 | **<0.001** | **-9.0** | **<0.001** |
| Bladder | 0.37 | 0.34 | 0.39 | 28.82 | **<0.001** | **-25.9** | **<0.001** |
| Grooming | 0.01 | 0.00 | 0.02 | 1.38 | 0.167 | -1.4 | 0.167 |
| Toilet use | 0.12 | 0.10 | 0.14 | 13.17 | **<0.001** | -12.8 | **<0.001** |
| Feeding | 0.17 | 0.15 | 0.18 | 20.25 | **<0.001** | **-19.4** | **<0.001** |
| Transfers | 0.11 | 0.10 | 0.13 | 12.28 | **<0.001** | -12.9 | **<0.001** |
| Mobility | -0.03 | -0.05 | -0.01 | -2.78 | 0.006 | **-2.6** | 0.009 |
| Dressing | 0.49 | 0.48 | 0.51 | 53.89 | **<0.001** | **-42.2** | **<0.001** |
| Stairs | 0.22 | 0.20 | 0.24 | 20.88 | **<0.001** | **-18.9** | **<0.001** |
| Bathing | 0.10 | 0.09 | 0.11 | 17.55 | **<0.001** | **-17.0** | **<0.001** |
| **Total BI** | **1.66** | **1.58** | **1.73** | **40.97** | **<0.001** | **-37.3** | **<0.001** |

The cross-tabulation table continues to tell a similar story to that described for the ratings in admission (see *Table 56*).

**Table 56: Cross tabulation of NPDS-BI and FIM-BI ratings for Urinary continence and Dressing on discharge**

| **Bladder** | **NPDS-BI** | | | |
|---|---|---|---|---|
| FIM-BI | 0 | 1 | 2 | **Total** |
| 0 | 668 | 80 | 92 | 840 |
| 1 | 134 | 64 | 53 | 251 |
| 2 | 829 | 246 | 2517 | 3592 |
| Total | 1631 | 390 | 2662 | 4683 |

| **Dressing** | **NPDS-BI** | | | |
|---|---|---|---|---|
| FIM-BI | 0 | 1 | 2 | **Total** |
| 0 | 631 | 7 | 3 | 641 |
| 1 | 1700 | 520 | 158 | 2378 |
| 2 | 160 | 462 | 1042 | 1664 |
| Total | 2491 | 989 | 1203 | 4683 |

## *Discussion and conclusion*

Conversion to a Barthel Index provides the opportunity to compare directly the level of independence as rated by the nursing staff and by the therapy team.

In this analysis the NPDS-BI tended to give lower BI scores than the FIM-BI. Although the gap between individual item ratings had narrowed somewhat by discharge, this was a consistent finding, most evident in the domains of 'Bladder' and 'Dressing'. At the time of admission, it is quite possible that scoring difference could be explained by one or other group of professionals simply not being as familiar with the patient's level of performance. However, by discharge the treating team will know the patient very well indeed, and the fact that the same pattern is seen, suggests that this is due to differences in the algorithm or to real differences in performance, rather than a lack of familiarity.

In the case of bladder function, the FIM does not include any means to identify the fact that a patient may be catheterised, and in this sense may give a less accurate BI translation than the NPDS - which does identify the presence of a catheter, as well as the frequency of accidents.
In the case of dressing, however, the differences were most likely accounted for by differences in the level of performance during therapy sessions and in day-to-day activities on the ward.

Patients with neurological disability may demonstrate variable performance depending on their motivation and levels of fatigue. The nursing staff, who see the patient all around the clock, may observe variations that are not evident during therapy sessions, when patients are likely to try their hardest. Thus the finding that therapists were more likely to perceive and rate the patient as being more independent is not unexpected, but it underlines the need for nurses to be involved in rating of functional outcomes, as well as the therapy team. In the meantime, when derived BI scores are available from both the FIM and the NPDS, we recommend using the NPDS-BI data for any outcomes analysis, as this is likely to provide an accurate translation of bladder function and to be more representative of the patient's round-the-clock level of function.

# Outcomes – Goal Attainment Scaling (GAS)

In recent years, increasing emphasis is placed on what the end-users of the NHS (i.e. patients and their families) actually want from their treatment. Measures that enable patients to record their outcomes and experience (PROMs and PREMs) are afforded increasing importance.

Patients undergoing neurorehabilitation present with a diverse pattern of impairments and disabilities. They have varied potential for improvement, and differing priorities for treatment. Whilst the UK FIM+FAM provides a useful yard-stick for measurement, covering a broad range of physical, cognitive, communicative and psychosocial function, no single standardised measure can capture the full range of patients' individual goals for rehabilitation. Moreover, changes recorded on the FIM+FAM do not necessarily reflect what the team set out to achieve, nor whether these met the expectations of the patient and/or their family, and matched their priorities.

Goal-setting has become a standard part of practice in rehabilitation.[158] Goal attainment scaling (GAS) is a method for assimilation of achievement in a number of individually-set goals into a single aggregated 'goal attainment score', providing a person-centred outcome, focused on that individual's priorities. Originally described by Kirusek and Sherman in the 1960s[159] it has been applied in various areas of complex intervention[158, 160, 161] including brain injury rehabilitation.[162]

GAS offers a number of potential advantages as an outcome measure for patients with complex disabilities. As well as providing a quantitative assessment of goal attainment, it also affords qualitative information about the patient's priority goals for treatment and their respective importance. The process of goal-setting and rating supports dialogue between the patient and their treating team, and offers an additional opportunity to negotiate mutually agreed expectations for outcome. However, clinicians require sufficient knowledge, training and experience to support patients to set realistic goals.[163]

The use of GAS is still somewhat controversial.

- Although large scale studies on inter-rater reliability are lacking, the small studies published to date are generally favourable.[164, 165] Some authors have been impressed by its responsiveness and sensitivity to patients' values[162, 166] and its flexibility across the domains of impairment, disability and participation.[161]

- Others (typically coming principally from a scientific measurement perspective, as opposed to a clinical perspective) have raised concerns about the mathematical concepts underlying the tool – particularly its non-linearity[163, 167, 168] and lack of uni-dimensionality.[169]
- Contrary to the originators' assertions that the GAS formula produces interval quality data, Steenbeek and colleagues[163, 168] point out that as GAS is based on a 5-point scale, the data can at best be only of ordinal quality and they recommend the use of medians and non-parametric statistics.
- In addition, there is a plethora of scoring systems in current use, each assigning different numerical scales for GAS which do not support assimilation of data.
- On the ground, clinicians raise the concern that the application off GAS as originally described by Kiresuk and Sherman,[159] is excessively time consuming for use in the routine clinical setting.[167]

Like it or hate it, GAS has a rapidly expanding literature[158] and there is growing interest from clinicians who, frustrated by the limitations of standardised scales, are starting to take a broader view of outcome assessment. Nevertheless, if it is to be used as part of the armamentarium for outcome measurement in rehabilitation, it must both be feasible to use in the context of routine clinical practice. Moreover the data should be interpretable and must be shown to provide added value.

Within this programme, we explored the following aspects of goal attainment scaling:

1. We developed a standardised approach to the application of GAS.[58]
2. We then examined the extent to which it provides added value over and above the other standardised measurement tools.[59]
3. Having established its added value, we explored way to make it practical for use in the clinical setting – the GAS light method.[58]
4. The impact of assigning different numerical scales.[170]

All of these aspects are described in more detail in the respective publications, but will be summarised here.

## *A standardised approach to GAS*

GAS builds on the fact that goal setting is a routine part of clinical practice in most clinical settings. As well as being an outcome measure, it supports

- **Communication and collaboration** and between the multi-disciplinary team members as they meet together for goal-setting and scoring.
- **Patient involvement** – GAS support both involvement and autonomy for patients and their families. There is emerging evidence that goals are more likely to be achieved if patients are involved in setting them. GAS is also shown to have positive therapeutic value in encouraging the patients to reach their goals.[171]
- **Establishing shared expectations -** the more formalised process of 'a priori' goal setting and defining and agreeing expected levels of achievement with the patient and their family supports the sharing of information at an early stage of rehabilitation and the negotiation of realistic goals for the programme.

We introduced a standardised seven-step GAS procedure which is outlined in Table 57 below:

**Table 57: The seven-step standardised procedure for using GAS in clinical practice**

| Seven Steps | Description |
|---|---|
| **Goal setting** | Out of the defined set of individual objectives for their programme, 1-6 priority 'personal goals' are identified and agreed between the patient (and/or family carer) and their treating team during the initial goal-planning meeting. |
| **Goal definition** | Every effort is made to maximise patient involvement in identifying their personal goals, which are then made 'SMART' (Specific, Measurable, Achievable, Realistic and Timed) through a process of negotiation with the patient and/or family |
| **Goal weighting (optional)** | The chosen goals may weighted by 'importance' (determined by the patient/family) and the degree of 'difficulty' (judged by the treating team). These are each graded on a scale of 0-3 ranging from 0='not at all' to 3='very' important or difficult,[172] and goal weighting is the multiplicand of 'importance x difficulty'. |
| **Baseline scores** | In order to allow for deterioration, and in accordance with previous applications,[172-174] baseline scores for each goal are allocated on admission as '–1' unless no clinically plausible worse outcome is possible - in which case a score of '–2' is given. (For example, for a patient whose goal is to be able to walk independently indoors with a walking aid: if at baseline they were starting to take some steps with the assistance of two people, this would score -1. However, if they were unable even to stand, let alone take steps, this would score -2). |
| **Goal evaluation** | At discharge from the programme, goal attainment is reviewed together with the patient and/or family, and rated on a five-point scale where:<br>'0' denotes the expected level of achievement,<br>'+1' and '+2' are respectively 'a little' and 'a lot' better than expected,<br>'-1' and '-2' are correspondingly 'a little' and 'a lot' less than the expected level.<br>Rating is based on their actual performance in relation to the expected level of achievement for each goal (i.e. what they do, not what they could do). In the case of disagreement, the lower level is scored. |
| **GAS** | The attainment levels for the chosen personal goals are then combined in a single aggregated 'T score' by applying the formula recommended by Kiresuk et al.[159] which accounts for variable numbers of goals, inter-correlation of goal areas and variable weighting:<br>Total score = $50 + \{[10\Sigma(w_i x_i)]/[0.7\Sigma w_i^2 + 0.3(\Sigma w_i)^2]^{½}\}$,<br>where $w_i$ = with assigned to the *i* th goal and $x_i$ = the score of the *i* th goal. |
| **Interpretation** | If goals are set in an unbiased fashion so that results exceed and fall short of expectations in roughly equal proportions, over a sufficiently large number of patients, one would expect a normal distribution of GAS T scores with a mean of 50 and standard deviation of +/-10. |

GAS therefore depends on two things – the patient's ability to achieve their goals and the clinician's ability to predict outcome, which requires knowledge and experience. Some professionals may find this challenging, but we believe that if a clinician is providing an intervention, they should have some idea about the likely outcome, and using GAS has helped us to develop our skills in outcome prediction. It is not necessary to be correct all of the time – so long as goals and over and under-achieved on a more or less equal basis. As noted above, the demonstration of a mean T score around 50 provides feedback relating to the accuracy of our goal-setting.

GAS is conceptually different from standardised measures – if interval measures may be described as measuring with 'a straight ruler', and ordinal measures as 'a piece of string', then GAS is the equivalent of measuring with a series of elastic bands. Many clinicians reared in the tradition of rigorous and objective measurement struggle with this concept.

Standardised measures still provide a useful yard stick for comparing different populations of patients on a level platform and it is NOT suggested that GAS should replace them. However, it does provide a useful reflection of outcomes that are of critical importance <u>to the patient in the context of their own lives</u>, which is something not provided by traditional measures. For this reason we recommend that GAS and standardised measures are used side by side.

## *Added value of GAS*

We examined the relationship between Goal Attainment Scaling (GAS) and three commonly-used global disability measures (the Barthel Index, the FIM and FIM+FAM) in the assessment of outcome from an in-patient rehabilitation programme. We explored both the quantitative aspects of measurement, and the qualitative nature of the goals set, to determine whether GAS has the potential to offer added value as a person-centred outcome measure for rehabilitation following brain injury. Our specific research questions were:

- What is the relationship between individualized goal attainment scaling (GAS) and the standard global outcome measures, and is GAS more responsive?
- What types of goals are commonly chosen by individuals undergoing rehabilitation, and to what extent do they overlap with standardised measures?

Goals were mapped onto the World Health Organisation's International Classification of Functioning Disability and Health (ICF).[119]

Once again as the data are already published,[59] only a summary is presented here.

## *Methods*

### Design and setting:

The study was undertaken in a tertiary regional specialist neuro-rehabilitation service for younger adult patients (predominantly 16-65 years) with complex neurological disabilities in the UK. In a prospective cohort analysis, routinely collected standardised outcome data (the Barthel Index and UK FIM+FAM) were compared with GAS scores for consecutive in-patients admitted to the unit for rehabilitation following acquired brain injury (of any cause) during a three-year period between 1st March 2005 and 28th February 2008.

### Participants:

A total of 164 brain-injured patients were admitted for neurorehabilitation during this period – 102 (62%) males and 62 (38%) females. Their mean age was 44.8 (SD 14.4) years, and the mean length of stay was 88 days (SD 51). The cause of brain injury was stroke in 108 (66%), trauma in 30 (18%) and the remaining 26 (16%) had other causes, such as anoxia, inflammation and tumour.

### Measurements

The routine rating of target (or 'goal') scores for the UK FIM+FAM provided an opportunity to apply the GAS formula to FIM+FAM scores and facilitates direct comparison of the tools as a measure of the achievement of both personal and FIM+FAM goals. Therefore, in addition to analysis of raw sum scores for the FIM+FAM, we applied the same principles of goal attainment scaling (GAS) to the analysis of FIM+FAM data.

- For each patient, each item of the FIM+FAM was retrospectively allocated a score on the five-point scale of -2 to +2 (according to the rules below), rated both on admission (baseline) and at discharge (achieved).
- At baseline: If the FIM+FAM item level was not expected to change (i.e. the admission and target levels were the same) a baseline score of 0 was allocated. If the target level was higher than

baseline, a score of -1 was allocated, unless the baseline FIM+FAM score was 1 (i.e. worst possible) in which case -2 was allocated.

- At discharge: If the target level was achieved, a GAS rating of 0 was allocated. If the target level was not reached, -1 was allocated, unless it had deteriorated from baseline, in which case -2 was allocated. Similarly if the discharge level exceeded the target level, +1 was allocated unless level seven (complete independence) was achieved, in which case +2 was allocated.

The scores for the 30 items were then combined using the GAS formula to derive GAS-transformed FIM and FIM+FAM scores (baseline and T scores for Motor and Cognitive sub-scores, in each case). This transformation could not be applied to the BI, as its three-level structure does not allow translation to five levels.

Because no specific information was available to weight the FIM+FAM goals, goal weights were all set at 1, for GAS transformation. Therefore, to facilitate comparison, both weighted and un-weighted GAS scores were computed for the personal goals.

We also undertook a qualitative analysis of the personal goals that were chosen, in order to identify the common goal areas. These were mapped retrospectively onto the FIM and/or FAM items and also onto domains of the WHO ICF[119] with reference to the linking rules published by Cieza[175] and with the assistance of the ICF Illustration library online (www.icfillustration.com). Second level categories (three-digit codes) were used as they are considered to provide the best trade-off between breadth and depth of coding.[176, 177] Two investigators (LTS and HW) coded the goal descriptions independently, then pooled and discussed their results to produce an agreed ICF code (or set of codes) for each personal goal. The principal codes were then assembled for each of the common goal areas.

## Data handling and analysis

In this analysis, the majority of instruments yielded ordinal data. Although the GAS formula should theoretically deliver normally distributed data, one-sample Kolmogorov-Smirnov tests revealed that the data in this set did not conform to normality. Non-parametric statistical tests were therefore used wherever possible. Wilcoxon signed rank tests were used to evaluate changes in GAS and standardised measures from baseline.

Current techniques for evaluation of responsiveness rely on the estimation of means and standard deviations. Responsiveness was compared using the Effect Size (mean change from baseline/SD baseline) and the Standardised Response Mean (SRM) (mean change from baseline / SD change). Effect sizes have been reported to over-estimate response if the distribution of baseline scores is narrow, as they tend to be with GAS,[178] and the SRM avoids this problem to a certain extent. As effect size and SRM rely on parametric assumptions, Wilcoxon z values were also given. Spearman rank correlations were used to examine associations between the various measures.

## Results

Descriptive statistics for the admission (baseline) and discharge (achieved) scores are shown in *Table 58*. Weighting made very little difference to the personal GAS scores, as the weighted and un-weighted (achieved) GAS T scores were highly correlated (Spearman rho 0.9, p<0.001) with no systematic bias between them (Wilcoxon z = -1.3, p=0.19).

All measures changed significantly from baseline to discharge (Wilcoxon z = -7.9 to -11.0; p<0.001). The effect sizes were all 'large', interpreted according to Cohen[71] (see foot note to Table 4.26), with the exception of the FIM and FIM+FAM cognitive scales. Median GAS (achieved) T scores were close to 50, both for personal goals and for GAS-transformed FIM±FAM data, indicating that goals exceeded and fell short of expectation in roughly equal proportions.

**Table 58: Summary of change from admission to discharge within the FIM, FIM+FAM and GAS**

| Measure | Admission Median (IQR) | Range | Discharge Median (IQR) | Range | Change score Median (IQR) | Wilcoxon z* | Effect size** | SRM |
|---|---|---|---|---|---|---|---|---|
| **Standardised measures: raw scores** | | | | | | | | |
| **Barthel Index** | 9 (4-13) | 0-20 | 16 (11-20) | 0-20 | 6 (2-9) | -10.4 | 1.0 | 1.37 |
| **FIM Motor** | 44 (27-65) | 13-91 | 73 (53-86) | 13-91 | 20 (9-32) | -10.6 | 0.93 | 1.44 |
| **FIM Cognitive** | 22 (16-28) | 5-35 | 27 (22-31) | 5-35 | 3 (0-6) | -9.3 | 0.47 | 0.94 |
| **FIM Total** | 67 (45-87) | 18-122 | 101 (77-115) | 18-126 | 24 (12-38) | -10.8 | 0.89 | 1.53 |
| **FIM+FAM Motor** | 53 (35-75) | 16-111 | 87 (66-103) | 16-112 | 26 (14-38) | -10.8 | 1.0 | 1.53 |
| **FIM+FAM Cognitive** | 61 (44-76) | 14-98 | 71 (58-80) | 13-91 | 6 (-1 – 15) | -7.9 | 0.36 | 0.75 |
| **FIM+FAM Total** | 118 (84-141) | 30-205 | 163 (129-184) | 30-209 | 36 (21-55) | -10.8 | 0.88 | 1.61 |
| **Goal attainment scores, including GAS transformed measures** | | | | | | | | |
| **Personal GAS (weighted)** | 35.0 (30.6-35.7) | 20-40 | 50.0 (44.2-51.8) | 24-64 | 14.4 (11.0-19.4) | -11.0 | 3.54 | 2.29 |
| **Personal GAS (unweighted)** | 35.0 (31.9 35.5) | 21-46 | 50.0 (46.4 - 50) | 25-65 | 14.5 (10.9-18.1) | -11.0 | 3.16 | 2.23 |
| **FIM GAS** | 35.7 (31.9-40.5) | 16-50 | 50.9 (47.1-54.8) | 27.1-73.9 | 15.3 (7.6-21.0) | -10.8 | 2.20 | 1.63 |
| **FIM+FAM GAS** | 36.2 (33.0-40.4) | 19-50 | 51.2 (47.8-54.7) | 28.3-75.2 | 14.1 (8.8-19.9) | -10.8 | 2.37 | 1.73 |

*For comparability with the un-weighted transformed scores, both weighted and un-weighted personal GAS scores are recorded*
*FIM = Function independence measure, FIM+FAM = UK Functional Assessment Measure, GAS = Goal Attainment Scaling.*
*SRM = Standardised Response mean = (Mean change from baseline/SD change). Effect size = (Mean Change from baseline/SD baseline)*
** Significance all p<0.001 **Effect sizes may be interpreted according to Cohen (0.2 = small, 0.5 = moderate, 0.8 = large)*

The effect sizes were substantially higher for both GAS and GAS-transformed scores than for the raw standardised measure scores but, for the reasons previously noted,[178] were probably over-estimated due to the small baseline SD. The standardised response means (SRM) were therefore thought to provide a better basis for comparison in this context. At 2.23, the SRM was substantially higher for personal GAS (un-weighted) than for the raw standardised measures, which ranged from 1.37 for the BI to 1.61 for the FIM+FAM. GAS transformation of the FIM+FAM scores improved the SRM only modestly to 1.73.

GAS (achieved) T scores were closely correlated with the change in GAS T score from baseline (rho 0.77, p<0.0001). For comparison with other measures, the aggregated GAS T-score was used in preference to the change in GAS score from baseline, as it is inherently a measure of change. The relationship between change in the standardised measures between admission and discharge and GAS T scores is shown in *Table 59.*

**Table 59: Spearman Correlations between personal GAS T scores and change from baseline in other measures and their GAS transformed counterparts**

| | GAS Achieved T-scores | |
|---|---|---|
| **Measure** | **Personal GAS (unweighted) rho** | **Personal GAS (weighted) rho** |
| **Barthel Index** | 0.36* | 0.38* |
| **FIM Motor** | 0.36* | 0.39* |
| **FIM Cognitive** | 0.10 | 0.09 |
| **FIM Total** | 0.35* | 0.36* |
| **FIM+FAM motor** | 0.41* | 0.43* |
| **FIM+FAM Cognitive** | 0.12 | 0.09 |
| **FIM+FAM total** | 0.37* | 0.38* |
| **GAS Transformed measures (Achieved T scores)** | | |
| **FIM+FAM GAS** | 0.46* | 0.49* |
| **FIM GAS** | 0.41* | 0.43* |

**Correlations were significant at p<0.001*

As expected there was strong correlation between changes in the FIM+FAM and BI (rho 0.84, p<0.001). In contrast, only moderate correlations were seen between personal GAS T scores and the standardised measures (0.36-0.43 for the raw change scores, and 0.41-0.49 for the GAS-transformed FIM±FAM scores), suggesting that GAS may indeed encompass areas of change not included in the FIM+FAM or BI. We therefore undertook an analysis of the actual individual goals that were set, mapping these on to domains of the WHO's ICF[119] – see *Table 60.*

Whilst the philosophy of the unit is to encourage goals to be set as far as possible in areas relating to activities and participation, a small proportion (approximately 10%) were necessarily process goals concerned with areas such as discharge planning and setting up care - for example in patients with low awareness states. A further 5% of goals were set in domains of 'body functions' (i.e. impairment-related goals), but the remaining 85% addressed activities and participation. The most popular areas for goal setting were mobility, self-care and communication, which accounted for approximately 65% of the total goals set. Approximately 20% addressed extended or community-based activities such as domestic tasks or recreation / leisure activities. Goal attainment was generally quite consistent across all these goal areas, with 67-76% of goals being achieved or over-achieved.

Mapping of personal goals onto the FIM+FAM items could not be precise as the SMART goal description often did not coincide with FIM+FAM level descriptors, and some goals were reflected in multiple items of the standardised scale (for example, six FIM+FAM items address different aspects of toileting and incontinence). In some instances, a goal might lie within the area of a FIM+FAM item (e.g. walking) but may lie outside the range (for example walking over longer distances than 50 metres). Goals were therefore considered individually to assess whether they were likely to have been reflected by changes in the FIM and/or FAM scores. In all, 315 (47%) goals addressed areas that could feasibly have been reflected by changes in the FIM, and 413 (62%) by the FIM+FAM.

**Table 60: Analysis of goal areas including coverage by the FIM±FAM and ICF domains**

| Goal Category | No set | No. (% ) achieved | FIM | FIM * +FAM | FIM+FAM Items** | Principal ICF Domain codes – second level |
|---|---|---|---|---|---|---|
| **BODY FUNCTIONS** | **37** | **(67%)** | | | | |
| **Mental function** | | | | | | |
| Low level: Awareness / interaction | 10 | 7 | ± | ± | 18,29 | b110, b164 (d335) |
| Higher level: Memory / orientation | 11 | 7 | - | ± | 27,28,30 | b114, b144, b140, b180 |
| **Emotion / behaviour** | 5 | 3 | - | ± | 23,24 | b152, b147 |
| **Senses:** Vision/hearing | 7 | 5 | - | - | - | b210, b230, b250 |
| Pain | 4 | 2 | - | - | - | b280 |
| **ACTIVITIES** | | | | | | |
| **Motor function/coordination e.g.** | **41** | **(76%)** | | | | |
| Improving control of upper/lower limb | 41 | 31 | - | - | - | d440, d445, b710, b735 |
| **Mobility** | **168** | **(76%)** | | | | |
| Low level – standing/transfers/postural | 47 | 30 | ± | ± | 10-12 | d410, d415, d420, d465 (d450) |
| Medium level - walking / stairs | 97 | 79 | + | + | 14,15 | d450, d455, d460, d465 |
| High level – outdoors / running | 24 | 19 | - | - | - | d450, d455, d460, d465 (d920) |
| **Self-care** | **134** | **(75%)** | | | | |
| General independence | 13 | 9 | + | + | 1-6 | d599 |
| Eating/drinking/nutrition | 27 | 22 | + | + | 1,7 | d550, d560, b510, b530 |
| Toileting / continence | 33 | 23 | + | + | 6,8,9,11 | d530, b525, b620 (d420, d450) |
| Washing/dressing/grooming | 61 | 46 | + | + | 2-5 | d510, d520, d540 |
| **Communication** | **93** | **(69%)** | | | | |
| Total communication ( with aids etc.) | 29 | 18 | + | + | 17, 22 | d330, d335, d360, d399, d730, d760 |
| Speech – talking / understanding | 43 | 32 | + | + | 17,18,21 | d330, d350, d360, d730, b310, b320, b126 |
| Reading / Writing | 21 | 14 | - | + | 19,20 | d140, d145, d166, d170, d325, d345 |

| Goal Category | No set | No. (% ) achieved | FIM | FIM * +FAM | FIM+FAM Items** | Principal ICF Domain codes – second level |
|---|---|---|---|---|---|---|
| **PARTICIPATION** | | | | | | |
| **Extended activities of daily living*** | 47 | (74%) | | | | |
| **Cooking/meal preparation** | 37 | 28 | - | - | (EADL) | d630, d620 |
| **Household/gardening** | 9 | 6 | - | - | (EADL) | d620, d640, d650, d920 |
| **Finance** | 1 | 1 | - | - | (EADL) | d860 |
| **Community access** | 21 | (71%) | | | | |
| **General** | 8 | 6 | - | + | 13, 16 | d460, d470, d499 (d620) |
| **Using public transport** | 10 | 6 | - | + | 16 | d460, d470, (b164) |
| **Driving** | 3 | 3 | - | - | - | d475 |
| **Recreation / leisure** | 31 | (68%) | | | | |
| **Computers / using EAT** | 11 | 9 | - | ± | 25 | d360, d920, d825, b140, b144 |
| **Active** | 5 | 2 | - | ± | 25 | d920, d455 |
| **Arts/crafts** | 3 | 2 | - | ± | 25 | d920, d650 |
| **Social and general** | 12 | 8 | - | ± | 22, 25 | d910, d920, d760 |
| **Work/ education / responsibility** | 28 | (68%) | | | | |
| **Work** | 16 | 10 | - | - | - | d825, d845, d840 |
| **Education** | 4 | 3 | - | - | - | d820, d830 |
| **Parenting** | 8 | 6 | - | - | - | d660, d760 |
| **OTHER GOALS** | | | | | | |
| **Process e.g.** | 67 | (86%) | | | | |
| **Discharge / care** | 61 | 54 | - | - | - | e115,e120, e125, e310, e340, e575. |
| **Spending time at home** | 6 | 4 | - | - | - | |
| **Total** | **667** | **495** | | | | |

*EAT- Electronic Assistive Technology, FIM=Function Independence Measure, FIM+FAM = UK Functional Assessment Measure.*
*EADL= extended activities of daily living.*
*'+' denotes most goals reflected by changes in FIM/FAM, '±' some goals reflected (less then half), '-' not reflected*
*** Numbers refer to the item no in the UK FIM+FAM. The EADL items form a separate module not included in the original 30 FIM+FAM items. These are available from the authors on request.*
*ICF domain codes in brackets indicate codes from different constructs that were commonly included in the goal statement (for example community access goals frequently specified the purpose of shopping (d620).*

**In summary,** we found GAS to be more responsive than the standard instruments. Although there was considerable overlap between the personal goals chosen and those represented in the standard instruments, GAS also recorded gains in other important areas not addressed by those tools. Whilst previous studies have compared GAS to the BI[174, 178] and the FIM,[160, 166] this is the first published comparison with the FIM+FAM.

Our analysis of goals confirmed that the FIM+FAM items covered a larger proportion of personal goals (approximately two-thirds) than the FIM items (approximately half). This suggests that, even though the additional FAM items confer little benefit in terms of the measurement properties of the scale,[133, 136] they do appear to capture additional qualitative outcome information in areas of importance to patients. Even so, the personal goals covered a much wider area of personal experience than any of the standardised scales.

The calculation of effect size and SRM is a parametric technique that necessarily involves the computation of mean change and standard deviations, but the very different nature of the datasets may limit comparison. Because the large majority of baseline GAS scores are -1, the baseline variation in GAS score is small and this may lead to over-estimation of the effect size.[178] For this reason we used the Standardised response mean (SRM). Even so, the possible range of GAS (mainly 40-60) is very different from the FIM+FAM (30-210) so that differences in SRM could simply reflect the range of data, rather than the responsiveness of the instrument per se.

In this study we took the novel approach of applying the GAS formula to transform the FIM+FAM data to a similar range, so that personal and standardised scores could be compared 'on a level playing field'. The difference between them is then more likely to reflect the achievement of gain itself, rather than the way it is calculated. Applying this method, the SRM was approximately 50% greater for the individual GAS than the FIM+FAM.

We recognise that there is still considerable debate about the validity of GAS. A full rehearsal of the various arguments for and against it is beyond the scope of this section, but the findings reported here may go some way towards clarification. Mackay and Somerville 1996 have raised concern about the validity of goal weighting.[167] Whilst weighting may provide useful qualitative information for clinical interpretation of variance in goal achievement, it can sometimes have a perverse effect – especially the weighting for goal difficulty.[58] In this analysis, the inclusion of goal weighting made little difference in numerical terms. For research purposes, it would therefore appear reasonable to exclude it from the formula and so eliminate one possible source of bias.

## Strengths and limitations:

The strengths of this study are that it represents a sizeable cohort of data collected in the course of real life clinical practice.

- The inclusion of patients with brain injuries of any cause extends the perspective and experience of the study group more widely than if selection was limited to one diagnostic group.
- The systematic process of goal setting and negotiation on our unit ensured that the priority goals chosen were important to the patients, and the expectation for outcome was mutually agreed between staff and patients and/or their family.

**Limitations include**:

- This was a single centre study which may limit the generalisability of our findings.
- The mapping of goals onto the FIM+FAM and ICF was undertaken retrospectively. Although ICF linking rules were applied as carefully as possible, in the absence of a clearly stated rationale for all the goals, coding accuracy cannot be fully guaranteed.

- Similarly it was not possible to be 100% certain whether goal attainment was actually reflected in FIM+FAM levels, especially when the goal crossed multiple items.

Prospective coding would overcome these problems, but in its current form the ICF is unwieldy for most busy clinicians to use. However, as common goals linked to ICF core-sets start to emerge in the future, it is possible that the production of tools such as localised decision trees embedded in electronic records may assist with coding, making it a more feasible option for use in the context of clinical practice.[179]

**In conclusion,** this study provided further evidence that GAS is a responsive measure which identifies the achievement of person-centred goals for rehabilitation, which may not be detected by commonly-used standardised measures. In addition to quantitative assessment, it provides useful qualitative information about the patient's priority goals. Whilst it cannot replace standardised measures, GAS provides useful added value as a measure of outcome from rehabilitation. Applied by our simplified method within the context of a goal-orientated rehabilitation programme, it provides a valuable adjunct to routine measurement for very little further investment of effort.

## *Making it practical for clinical use – the GAS light method*

Feedback from clinicians in the early stages of the programme identified a number of key barriers to implementing GAS in their routine clinical practice.

1. **"It takes too long".**
   a. The process of negotiating goals can be time-consuming, but clinicians generally accepted this as a normal part of treatment.
   b. What they really objected to was the 'outcome guide' recommended by the originators which pre-specified descriptors for each of the five potential outcome levels.
2. **"The numbering system is dispiriting"**
   a. Clinicians disliked the fact that, after all the work to achieve the goal, you still scored "0".
   b. Moreover, if the patient made good progress but narrowly missed the full specification for goal attainment, they still scored '-1' (i.e. the same as baseline).
   c. Clinicians preferred to define goal attainment in words.

The GAS light method, reduces the number of steps involved:

1. GAS is <u>not</u> applied to every staged goal, but just to the 3-4 key objectives that are agreed as the most important to the patient.
2. Weighting is reduced to 'importance only" (or no weighting at all)
3. Instead of an 'outcome guide' defining each of the five outcome score levels (-2, -1, 0, +1 and +2), the team concentrates on defining very carefully the expected 'level 0' outcome at baseline.
4. Then, at the end of the programme, the team and patient agree the level of attainment based on the follow verbal rating scale which resonates with clinical decision-making:
   a. Was the goal achieved – yes or no
      i. If yes, was it 'as expected', a little more' or 'a lot more'?
      ii. If no, was it partially achieved, was the patient the same as at baseline, or were they worse.
5. This verbal rating scale however, has six (rather than five points). It is converted to the five-point scale as shown in *Table 61* below.

**Table 61: The GAS-light verbal scoring system with numerical conversion**

| | | | | Computerisation | |
|---|---|---|---|---|---|
| **At Baseline** | With respect to this goal do they have? | Some function | ❑ | -1 | |
| | | No function (as bad as they could be) | ❑ | | -2 |
| | | | | | |
| **At Outcome:** | **Yes** | A lot more | ❑ | +2 | +2 |
| | | A little more | ❑ | +1 | +1 |
| | | **As expected** | ❑ | **0** | **0** |
| **Was the goal achieved?** | **No** | Partially achieved | ❑ | (-1) | -1 |
| | | No change | ❑ | -1 | -2 |
| | | Got worse | ❑ | -2 | |

### Implementation:

Both the conversion of the verbal rating to numbers and the algorithm to derive the GAS T-score were built in to the UKROC software, order to keep clinicians away from numbers.

### The impact of assigning different numerical scales

This five-point scoring method is the most widely used in the research literature,[172-174] but on a clinical level it is counter-intuitive and this may partly explain its limited uptake by clinicians who, as noted above, prefer to evaluate goals on a 6-point verbal rating scale.

To address this problem, Steenbeek and colleagues[163, 168] have proposed a 6-point numerical model in which all baseline scores start at '-2'; '-1' denotes partial achievement and '-3' denotes worsening. The range of -3 to +2 means that the symmetry of the original scale is lost when applying the GAS formula. However, they argue that the data can only be of ordinal quality and so they recommend the use of non-parametric statistics in any event.

In the course of this programme we explored three different GAS scoring methods:

- Model A: Steenbeek's 6-point scale ranging from -3 to +2
- Model B: and alternative 6-point model in which '-0.5' denoted partial goal achievement within the score range of -2 to +2.
- Model C: the traditional 5-point model ranging from -2 to +2 as described by Kiresuk and Sherman[159]

In a head-to-head comparison, we used the same large dataset to compare the impact of three different methods of goal scoring on the calculation of goal attainment scales.[170] The full results are not repeated here, but may be summarised as follows:

Although the median achieved T-scores were broadly similar for all three methods, Models A and B produced systematic bias in different directions - Model B marginally over-estimated goal attainment in comparison with standard goal scoring (Model C), whilst Model A under-estimated it. Because baseline scores were markedly lower in Model A, change scores were significantly higher –in fact almost twice that of the other two scoring models.

Overall, Model B provided a closer fit to the standard rating system than Model A, maintaining a similar score range, whilst still discriminating partial goal achievement from ‘ no change’. A potential advantage of this system is that, by counting all ‘-0.5’ scores as ‘-1’, Model B rating can easily be converted back to standard rating for the purposes of comparison with other datasets which use the standard five-point system.

As Kirseuk points out, because change over time is built into the way in which GAS scores are derived, the T-score is in itself a measure of change and therefore obviates the need for measuring change from baseline.[180] However, many researchers still continue to record GAS change,[172-174] especially when comparing with change in other measures. It is therefore important to be aware of the impact of the rating system on GAS scores, if baseline and or change scores are used. Providing that only the achieved T-score is used, and non-parametric statistical methods are applied (as Steenbeek et al recommend)[168] Model A rating may still offer a measure of goal attainment which is roughly equivalent to the standard method.

**In conclusion**, the use of different goal rating methods may have significant impact on the results of goal attainment scaling. The addition of a ‘-0.5’ scoring option may allow teams to record partial achievement of goals which start with a baseline score of -1, whilst still maintaining the standard score range of -2 to +2. This may be especially advantageous using the GAS clinically. However, when used for scientific reasons we recommend that ‘-0.5’ scores are transformed to ‘-1’ in order to maintain parity with the standard rating systems

Further information about GAS and practical tools to support its implementation may be found on our website http://www.csi.kcl.ac.uk/gas-tool.html

# Outcomes for cognitive behavioural rehabilitation settings

From an early stage in the rehabilitation programme, it was recognised that the UK FIM+FAM might not be sufficiently sensitive as an outcome measure to detect meaningful change in the patients undergoing cognitive behavioural rehabilitation – particularly in more community based settings. Following a series of workshops with providers of such services in 2011/12, a programme of work was set up to collect parallel data on four measures:

- The UK FIM+FAM,
- The Mayo Portland Adaptability Inventory, (MPAI)[181]
- The St Andrew's Swansea Neurobehavioural Outcome Scale (SASNOS)[182]
- The Health of the Nation Outcome Scale (HoNOS) (Royal College of Psychiatrists)[183]

In the event, that project has been taken forward outside of this programme by the a group led by Prof Nick Alderman through Independent Neurorehabilitation Providers’ Association. It will be reported separately when the programme is complete.

In the meantime, the MPAI has been emerging on the ground as a tool that appears to complement and extend the areas covered by UK FIM+FAM – especially in relation to items such as family relationships, residential independence and substance abuse. The MPAI consists of three subscales – Abilities (12 items), Adjustment (8 items) and Participation (8 items). Each is rated on a scale of 0 (no problem) to 4 (severe problem). The overall scores are then combined (with re-scoring of 4 items) and converted to a set of T scores for comparison with the data from a US cohort.

In a preliminary evaluation, we worked with one of our Australian Collaborators to provide a comparison of the UK FIM+FAM and the MPAI in a small cohort of patients admitted one of their residential units. The following analysis is presented with the permission of the Governing Board of Brightwater Care Group and with the assistance of Kristylee Sharp and Karla Seaman.

We wished to determine whether the MPAI provide a more sensitive measure than the UK FIM+FAM for their client group.

## *Setting*

The Brightwater Oats Street facility in Perth, Australia provides a community-based residential rehabilitation service for people aged over 18 years, with an Acquired Brain Injury and issues of cognitive impairment. The 43-bed unit is made up of 9 houses offering gradually reducing levels of care and support ranging from full assistance and 24 hour care (Houses 1 and 2) to sheltered independent living (Houses 8 and 9). Patients entering the programme are placed in the correct house according to their needs for care and support, and then gradually progress to less supportive houses as they progress towards independence.

## *Outcome measures*

The UK FIM+FAM is recorded routinely on admission and discharge, and at yearly intervals in between. The MPAI is recorded every 3 months throughout their stay. In this snapshot of data we compared the item, subscale and total scores of the MPAI and UK FIM+FAM at admission, and either discharge or one-year reviewfor those patients who were still in the programme

## *Participants*

Data were collated for a consecutive cohort of 32 patients with acquire brain injury admitted to the service. Their average age was 43.9 (SD 12.8) on admission, Male:female ratio 3:1 years; Aetiology: 44% stroke, 41% traumatic and 15% other causes). The mean length of stay was 82.5 (sd 41) days with a range of 4 months to 3 years.

Descriptive statistics are set out in Table 62 for the FIM, FIM+FAM and EADL items, and in Table 63 for the MPAI and its domains. Table 64 shows the change scores for the two scales.

**Table 62: FIM scores at admission and review/discharge (N=32)**

| | Range | Mean | SD | Median | IQR |
|---|---|---|---|---|---|
| **Time from admission to review/discharge (weeks)** | 16-76 | 53 | (13.2) | 55 | (51, 60) |
| **Admission** | | | | | |
| **FIM+FAM motor score** | 16-112 | 72 | (30.2) | 74 | (58, 97) |
| **FIM+FAM cognitive score** | 14-86 | 53.5 | (20.1) | 59 | (41, 69) |
| **FIM+FAM total score** | 30-198 | 126 | (47) | 135 | (97, 160) |
| **Review / discharge** | | | | | |
| **FIM+FAM motor score** | 16-112 | 81.2 | (28.8) | 90.5 | (68, 104) |
| **FIM+FAM cognitive score** | 14-88 | 59 | (19.0) | 61 | (47, 73) |
| **FIM+FAM total score** | 30-197 | 140 | (44.9) | 147 | (122, 173) |
| **Total EADL** | **6-37** | **13.9** | **(8.1)** | **11** | **(7, 18)** |

**Table 63:** MPAI scores at admission and review/discharge (N=32)

| | Range | Mean | SD | Median | IQR |
|---|---|---|---|---|---|
| Time from MPAI assessment 1 to 2 (weeks) | 14-59 | 43 | (12.7) | 46 | (37, 52) |
| **Ability Subscale** | | | | | |
| **Time point 1** | | | | | |
| MPAI Ability T-score | 34-81 | 56 | (10.7) | 55 | (49, 61) |
| MPAI Ability rank | 1-4 | 3 | (0.9) | 3 | (2, 4) |
| **Time point 2** | | | | | |
| MPAI Ability T-score | 31-88 | 52.3 | (10.8) | 52 | (45, 59) |
| MPAI Ability rank | 1-4 | 2.8 | (0.8) | 3 | (2, 3) |
| **Adjustment Subscale** | | | | | |
| **Time point 1** | | | | | |
| MPAI Adjustment T-score | 43-71 | 54 | (7.3) | 54 | (48, 59) |
| MPAI Adjustment rank | 2-4 | 2.9 | (0.7) | 3 | (2, 3) |
| **Time point 2** | | | | | |
| MPAI Adjustment T-score | 36-65 | 52 | (7.0) | 52 | (46, 58) |
| MPAI Adjustment rank | 1-4 | 2.8 | (0.8) | 3 | (2, 3) |
| **Participation Subscale** | | | | | |
| **Time point 1** | | | | | |
| MPAI Participation T-score | 42-74 | 58.2 | (9.9) | 58 | (50, 65) |
| MPAI Participation rank | 2-4 | 3.1 | (0.8) | 3 | (2.5, 4) |
| **Time point 2** | | | | | |
| MPAI Participation T-score | 37-74 | 52.7 | (10.5) | 52.5 | (43.5, 59) |
| MPAI Participation rank | 1-4 | 2.7 | (0.9) | 3 | (2, 3) |
| **Total MPAI score** | | | | | |
| **Time point 1** | | | | | |
| Total MPAI T-score | 38-81 | 56.3 | (9.9) | 54 | (50-63) |
| Total MPAI outcome rank | 1-4 | 3.1 | (0.9) | 3 | (2.5, 4) |
| **Time point 2** | | | | | |
| Total MPAI T-score | 34-75 | 52.4 | (9.7) | 52.5 | (44.5, 59.5) |
| Total MPAI outcome rank | 1-4 | 2.8 | (0.95) | 3 | (2, 3.5) |

**Table 64:** Changes in FIM, FIM+FAM and MPAI scores between admission and review/discharge

| Outcome measure | Mean change | 95% CI Lower | 95% CI Upper | t | p |
|---|---|---|---|---|---|
| **FIM (N=32)** | | | | | |
| Motor sub-score | 6.5 | 2.9 | 10.1 | 3.71 | <0.001 |
| Cognitive sub-score | 1.9 | 0.1 | 3.8 | 2.15 | <0.05 |
| **Total FIM** | 8.5 | 3.9 | 13.0 | 3.82 | <0.001 |
| **FIM+FAM (N=32)** | | | | | |
| Motor sub-score | 9.0 | 4.7 | 13.4 | 4.21 | <0.001 |
| Cognitive sub-score | 5.4 | 1.0 | 9.8 | 2.51 | <0.01 |
| **Total FIM+FAM** | 14.4 | 6.7 | 22.2 | 3.79 | <0.001 |
| EADL (N=29) | | | | | |
| EADL new | 5.0 | 2.2 | 7.8 | 3.70 | <0.001 |
| **MPAI (N=32)** | | | | | |
| Ability | -3.3 | -5.6 | -1.0 | -3.0 | <0.01 |
| Adjustment | -2.3 | -4.7 | 0.1 | -2.0 | NS |
| Participation | -5.5 | -7.7 | -3.3 | -5.1 | <0.001 |
| **Total MPAI** | -3.9 | -6.3 | -1.6 | -3.4 | <0.01 |

Figures 33 and 34 illustrate the changes in median item scores for the two scales. (In order to make these comparable, the score order has been reversed for the MPAI.)

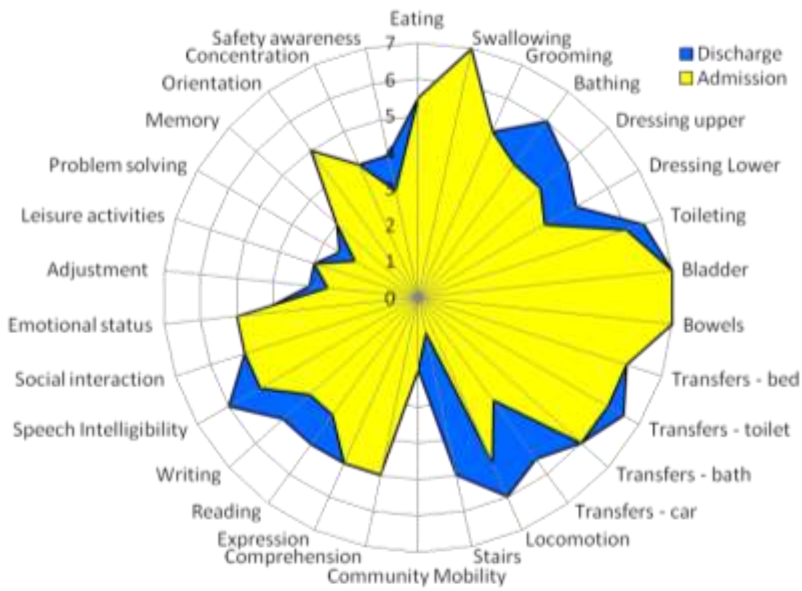



Figure 33: **Change in median FIM+FAM scores from admission to review/discharge**

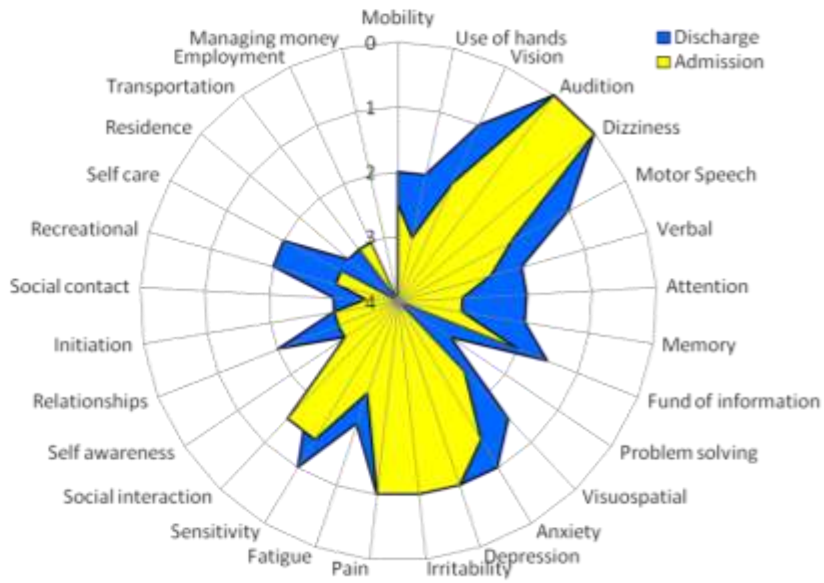



Figure 34: **Change in median MPAI scores from admission to review/discharge**

Taking this sample as a whole, both instruments detected clinically significant change between the two time-points, and changes were observed in most (but not all) of the items.

On the basis of this analysis, the MPAI had little obvious advantage over the FIM+FAM in terms of sensitivity although it did address items not covered by the FIM+FAM. Further analysis of change in the individual house 'stages' is currently underway to see if it provides better discrimination at the highter-functioning end of the programme.

# Outcomes – other tools

Within any national database, completeness of data is a critical issue, so it is necessary to constrain the number of measures used to a manageable core dataset. The instruments described so far within this section form that core toolset in UKROC.

However, the programme is wide-reaching and we anticipate that the range of tools will expand in due course as the scope of the programme becomes more broad.

A number of other tools have been explored, either as part of this programme or in parallel related NIHR-funded programmes, which are worthy of a brief mention here. These include two tools for use in community rehabilitation settings:

**The Workability Scale**[152, 153] is a new measure that was developed in the course of this programme as part of a wider collaboration with colleagues at the Auckland University of Technology, New Zealand. It is designed to:

- Assess the individual's ability to work and support needs in the context of their normal work environment, following the onset of acquired disability,
- Support decision-making with regard to vocational rehabilitation including withdrawal from work where appropriate.

It encompasses the complexity of physical, cognitive and behavioural challenges that are typically associated with neurological disability. However, it also has application in the more general context of work-related disability. Papers describing its conceptualisation, development and evaluation have been published during this programme.[114,115]

**The Needs and Provision Complexity Scale.**[184, 185] The original conceptualisation of the NPCS was as a community version of the Rehabilitation Complexity Scale. It is a brief, pragmatic tool to measure both the *needs* that an individual has for rehabilitation and support - and the extent to which those needs are *met* through service provision. It may be used at an individual level to monitor the changing needs of a given patient over time and the services that are provided to support them at different stages along the care pathway. Papers describing its evaluation and application were published as part of this programme in conjunction with a parallel project on community rehabilitation funded through the NIHR HS&DR programme.[184, 185]

As noted above these tools were primarily relevant for use in community settings, which was not the primary focus of this stage of the programme. However, they are likely to come into their own in the next stage of work when we plan to take a similar approach to the development of a national database for monitoring rehabilitation needs, inputs and outcomes in community-based rehabilitation programmes.

## *Summary of Chapter 4*

Within this chapter we have achieved our key deliverables, which were a set of psychometrically robust tools to measure needs inputs and outcomes that are fit for incorporation into the UKROC database, adapted and customised where necessary for the various settings in which they will be used.

The iterative process of development, evaluation and psychometric testing of the various tools that make up the UKROC database has continued throughout his programme. As a result of this work we have established a robust set of measures for needs, inputs and outcomes from rehabilitation that, for the most part, are demonstrated to meet the standards laid down by the Medical Outcomes Trust. Their reliability has been established and an understanding of their factor structure and scaling properties has underpinned the approach to the data analyses presented later in this report.

## Changes from the application

The purpose of this programme was to develop the tools that would be used to inform case-mix and tariffs for rehabilitation under the Payment by Results programme. Its direction has therefore been highly dependent on the direction and central prioritisation of tariff development in this area.

The major focus for currency and tariff development under PbR and subsequently NHS England has been within the Level 1 and 2 specialist inpatient rehabilitation services. The tools for these services are now fully established and we have continued to publish psychometric evaluations as the datasets have reached a sufficient sample size. Key tasks for the final stages of this programme were finalising the tools and implementing the UKROC database as the commissioning dataset for NHS England in parallel with the new commissioning arrangements. This has been achieved.

It was originally expected that tariff development for Level 3 and community-based services would progress in parallel with the specialist services and that we would work with the PbR team to develop similar tools for these areas as well. As a result of re-structuring, however, tariff development was halted in these areas and has not been taken forward.

Instead, in 2011-13, the need for tariffs for community-based residential rehabilitation for patients with complex needs in slow-stream and specialist nursing home settings was flagged and a suitable set of tools (including the RCS-SNH, an adapted version of the Rehabilitation Complexity Scale for use in specialist nursing homes, and the Northwick Park Dependency Scale) has been developed. We also worked with a group of neuro-behavioural rehabilitation providers to determine whether any additional outcome measures should be included in the UKROC database.

However, a proposal from the Clinical Reference Group for Specialist Rehabilitation (CRG-SR) to NHS England to include this slower stream rehabilitation within the service specification was not progressed due to the current heavy financial pressures on NHS England which have tended to reduce, rather than expand, the scope of the existing directly prescribed service specifications.

In the absence of this central commissioning with its drive for national benchmarking of these services, it is unclear whether this dataset will actually be implemented. Therefore, although the toolset has been developed, there are currently no plans to implement it within the UKROC dataset until there is clearer evidence about how ongoing data collection would be commissioned for this area of practice in the future. The tools are, however, made freely available through our website for those providers who wish to use them on a voluntary basis. In our view, this work has therefore progressed as far as it can in the lifetime of this research programme.

### Directions for further research and development

Within this first seven years, the primary focus has been on neurorehabilitation, but as the scope of data collection broadens to encompass different types of service, new tools (or adaptations of the existing ones) will continue to be added.

# Chapter 5: Current provision of rehabilitation resources

## Objective:

To apply the tools in a variety of neurorehabilitation settings and to identify the rehabilitation resources currently provided in relation to caseload complexity

### *Key deliverables*

1. Analysis of staff time in relation to complexity of rehabilitation needs to inform the weighted payment model.
2. Analysis of caseload complexity in relation to resources (medical, nursing and therapy staff) to inform designation of specialist rehabilitation services by NHS England.
3. A system for identifying patients with category A needs (requiring specialist (Level1/2a) services.

### *Background*

The Northwick Park nursing and therapy Dependency Scales (NPDS/NPCNA and NPTDA) both translate into staff hours, providing a crude measure of staff inputs. As discussed in Chapter 4, the NPCNA formula (which was designed originally for computation of care hours in the community) over-estimated time required for <u>direct hands-on</u> patient care in a hospital setting. Nevertheless, once the indirect and non-clinical activities of nursing and care staff are taken into account, it provides an estimate of overall staff time requirements that is adequate for comparison at least on a proportionate basis.

By applying the tools developed in Part 1 systematically across the different service settings, we are able to describe and measure the needs for care and the rehabilitation interventions offered in each service at any one time, and thus determine the relative complexity of the caseload. We also examined the relative resource implications (in terms staffing levels ) of providing rehabilitation for patients in different bands of complexity.

- We also expected to find that more complex patients receive higher levels of rehabilitation intervention, making more demands on staff time, and so requiring higher staffing levels in proportion with caseload complexity.
- Although each service will have a mix of simpler and more complex patients within their caseload, we expected to find Level 1 services carry a higher proportion of complex cases than Level 2 services.
- It was further recognised that Level 1 services for a heterogeneous group. Some cater for patients with primarily physical disabilities, whilst others take ambulant patients with challenging behaviours. Staffing requirements are clearly different for these patient groups.

By comparing the needs of the caseload with total staff time available from the rehabilitation resources provided, we could identify the staff establishment required to match a given caseload and so start to develop a generic formula for calculating staff requirements in relation to case-mix within the different levels of service.

### *Relevant policy development*

From 2013, patients with category A needs in Level 1 and 2a services were directly commissioned by NHS England. The Service Specification set out criteria for the designation of eligible services, based on complexity of their caseload, with reference to the BSRM Standards, see *Appendix 2.1*.

Level 1 services were expected to have higher staffing levels than Level 2a and b services to meet the demands of their more complex caseload, but the differences had not been quantified and neither had their respective cost implications.

NHS England relied on information supplied by UKROC to identify those services that were eligible for designation as a level 1/2a services. Services were signposted to the different services levels (i.e. Level 1 a, b, c or Level 2 a, b) on the basis of a) the complexity of their caseload and b) having the requisite staffing levels to manage that caseload.

UKROC data were used to set quantifiable standards for staffing provision within each service level and subsequently to calculate their differential costs. This analysis led to the development of a complexity-weighted costing model that was subsequently used to inform the development of tariffs for specialist rehabilitation services (see Chapter 7).

# Analysis of staff time in relation to complexity of rehabilitation needs to inform the weighted payment model.

In this analysis we set out to determine the relative proportion of rehabilitation inputs (nursing and therapy staff hours) provided for patients with different levels of rehabilitation needs as measured by the Rehabilitation Complexity Scale, broken down into five bands.

This analysis was conducted in two phases:

- Phase 1 was conducted in 2009 on data from a single centre as an exemplar to provide the initial modelling on which the principles of weighted bed day model were developed.
- Phase 2 was conducted in 2012 on a much larger multi-centre dataset to apply the model and develop the actual indicative tariffs that were published in the PbR guidance for 2013/14.

### *Phase 1 – initial modelling – single centre analysis*

The original model was developed on data from a single tertiary (Level 1) specialist inpatient neurorehabilitation service. During the 30-month period between June 2006 and December 2008, 1200 parallel ratings of complexity (RCS v8) scores, and dependency scores (NPDS and NPTDA) scores were collated for the consecutive cohort of all admitted patients - a total of 179 cases.

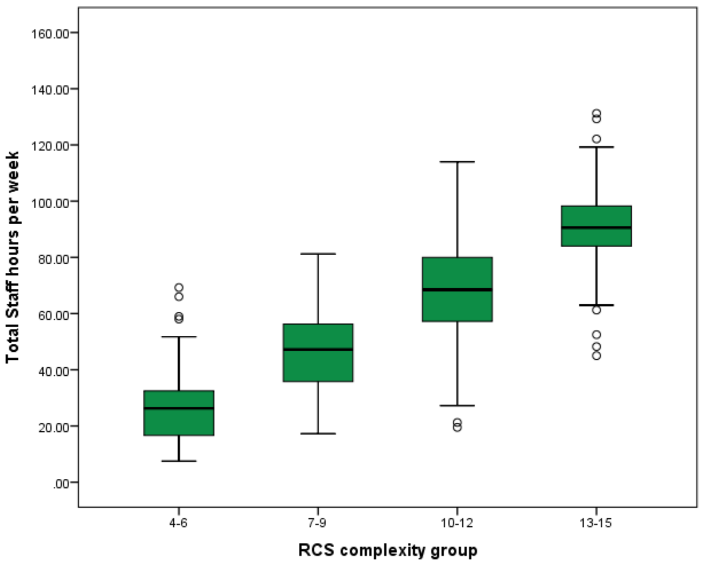
The first step in development of the weighted costing model was to determine the relative proportions of staff time associated with each of the five complexity bands. Each RCS score provides a snapshot of the case complexity of an individual for the week in which it was recorded, and the corresponding NPDS and NPTDA provide an estimate of the hours of therapy and nursing staff time that were utilised for that patient during that week.

*Table 65* shows a breakdown of the mean nursing and therapy hours per week for each of the five complexity bands from our 1200 parallel ratings. In this sample there were no RCS ratings in the 0-3 score range. The box plots in *Figure 36* confirms that the data for total staff hours/week within the remaining bands are normally distributed and separate quite cleanly into the four bands.

**Table 65: The distribution of nursing and therapy staff hours across the five complexity bands**

| Complexity band (RCS score) | Total no of ratings | Nursing/ care hours* | Therapy hours** | Total staff hours*** | Staff time ratio |
|---|---|---|---|---|---|
| | | Mean (SD) | Mean (SD) | Mean (SD) | |
| **Very High (13-15)** | 262 | 59.9 (8.6) | 32.2 (19.9) | 90.5 (13.0) | **1.9** |
| **High (10-12)** | 558 | 45.7 (15.7) | 26.7 (6.6) | 68.4 (16.2) | **1.5** |
| **Medium (7-9)** | 321 | 27.7 (13.7) | 18.8 (5.3) | 46.5 (13.7) | **1.0** |
| **Low (4-6)** | 59 | 11.5 (11.1) | 15.9 (5.7) | 27.4 (13.7) | **0.6** |
| **Very Low (0-3)** | 0 | - | - | - | - |
| **Whole sample** | **1200** | **41.6 (19.4)** | **22.8 (8.1)** | **64.8 ( 25.1)** | |

*RCS = Rehabilitation Complexity Score*
**Nursing care hours estimated by the computerised algorithm in the Northwick Park Care Needs Assessment*
***Therapy hours estimated by the computerised algorithm in the Northwick Park Therapy dependency Assessment.*
****Total staff hours are the sum of the estimated nursing and therapy hours for each patient*




**Total staff hours per week are estimated through the Northwick Park nursing and therapy Dependency Scales*

**Figure 35: The distribution of total nursing and therapy staff hours per week* across the different complexity bands**

## *Phase 2 – Multi-centre modelling*

The original model was developed on data from a single tertiary (Level 1) specialist inpatient neuro-rehabilitation service. It required validation on a multicentre sample. We also wished to determine whether the banding factor, based on proportion of staff time within the different complexity groups, was the same or different in the Level 1 and 2 services.

By November 2012, the UKROC dataset had recorded a total of 33,230 serial RCS-E scores from a total of 70 different services (RCS-E version 12 was the current version in use at that time). According to the service level classification at the time, there were 13815 ratings from a total of 17 Level 1 services, 13752 from 20 Level 2 services and 5663 ratings from a total of 33 services, which had only recently started to submit data and were not yet signposted.

The UKROC software automatically records and converts item-level RCS-E scores to the standard RCS scores (RCS version 8) to enable comparison between the two scores. *Table 66* shows the conversion of total RCS to RCS-E-v12 scores. The cut-off scores for the complexity bands in this analysis were: Very Low = 0-4, Low = 5-7, Medium= 8-10, High = 11-14, Very high = 15-20. These cut-offs were the same for the High and Very High bands as in our original analysis, but one point different in the lower three bands (see Table 12, page 45 in Chapter 4). We used the new cut-off points in the analysis of RCS-E and staff hours below.

**Table 66: Conversion of RCS to RCS-E-v12 scores (translation table)**

| RCS-v8 | Valid | Missing | Mean | Std. Dev. | Minimum | Maximum | Median | Percentiles (IQR) | |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 94 | 3 | 0.45 | 0.541 | 0 | 2 | 0 | 0 | 1 |
| 1 | 148 | 10 | 1.59 | 0.581 | 1 | 3 | 2 | 1 | 2 |
| 2 | 306 | 21 | 2.63 | 0.571 | 2 | 4 | 3 | 2 | 3 |
| 3 | 546 | 47 | 3.67 | 0.617 | 3 | 5 | 4 | 3 | 4 |
| 4 | 856 | 48 | 4.66 | 0.588 | 4 | 7 | 5 | 4 | 5 |
| 5 | 1040 | 53 | 5.75 | 0.658 | 5 | 8 | 6 | 5 | 6 |
| 6 | 1369 | 56 | 6.81 | 0.681 | 6 | 9 | 7 | 6 | 7 |
| 7 | 2252 | 66 | 7.86 | 0.691 | 7 | 11 | 8 | 7 | 8 |
| 8 | 3332 | 83 | 8.98 | 0.698 | 8 | 12 | 9 | 9 | 9 |
| 9 | 4543 | 124 | 10.14 | 0.741 | 9 | 13 | 10 | 10 | 11 |
| 10 | 5333 | 146 | 11.39 | 0.825 | 10 | 14 | 11 | 11 | 12 |
| 11 | 5601 | 127 | 12.69 | 0.891 | 11 | 15 | 13 | 12 | 13 |
| 12 | 4285 | 79 | 14.03 | 0.894 | 12 | 17 | 14 | 13 | 15 |
| 13 | 2334 | 40 | 15.47 | 0.959 | 13 | 18 | 15 | 15 | 16 |
| 14 | 932 | 31 | 16.68 | 0.989 | 14 | 19 | 17 | 16 | 17 |
| 15 | 259 | 16 | 17.72 | 1,043 | 15 | 20 | 18 | 17 | 18 |
| | **33230** | **950** | | | | | | | |
| | | **34180** | | | | | | | |

A total of 5868 parallel scores for RCS-E, NPDS and NPTDA were available from:

- 5 Level 1a services (n=1154 parallel pairs),
- 7 Level 1b services (n=968 parallel pairs),
- 9 Level 2a services (n=1218 parallel pairs),
- 13 Level 2b services (n=1751 parallel pairs),
- (12 other services which had yet to be classified (n=777 parallel pairs)).

The distribution of therapy and nursing hours within each RCS-E band are shown in *Table 67*, and illustrated in *Figure 36*.

**Table 67: Total staff per week by RCS-E-v12 band**

| RCS-E v 12 | | | Nursing hours | | Therapy hours | | Total Staff Hours | | |
|---|---|---|---|---|---|---|---|---|---|
| Band | n | % | Median | IQR | Median | IQR | Median | IQR | Ratio |
| Very High | 946 | 16% | **61.3** | 52.8-68.3 | **22.3** | 16.8-28.3 | **84.3** | 72.5-92.5 | **2.0** |
| High | 2912 | 50% | **45.5** | 31.5-63.0 | **17.5** | 12.3-23.5 | **64.9** | 47.9-81.2 | **1.5** |
| Medium | 1403 | 24% | **26.3** | 14.0-38.5 | **14.3** | 9.9-19.9 | **41.8** | 28.0-57.2 | **1.0** |
| Low | 490 | 8% | **14.0** | 7.0-24.5 | **12.4** | 8.2-17.3 | **27.3** | 18.2-40.0 | **0.6** |
| Very Low | 117 | 2% | **10.5** | 3.5-15.8 | **10.5** | 6.7-15.8 | **23.0** | 13.3-33.2 | **0.5** |

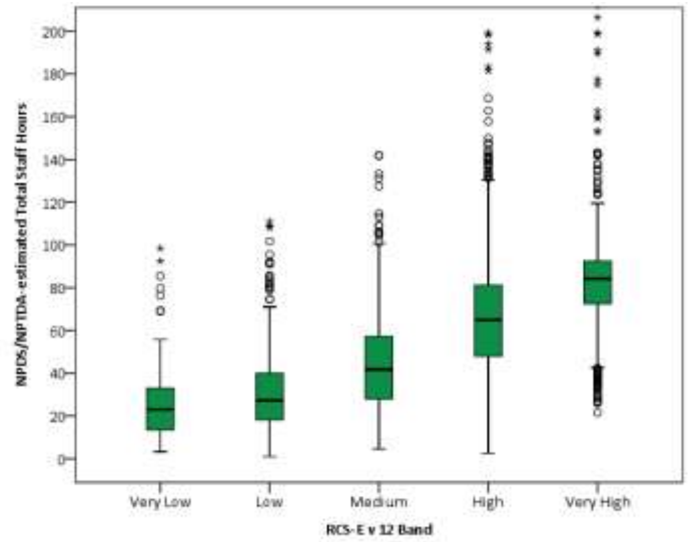



**Figure 36: Total staff hours per week by RCS-E-v12 band**

As expected, the total staff hours were slightly higher for the Level 1 services compared with the Level 2 services – see *Figure 37*, and with a somewhat different shape of the curve. Based on these differences, the staff time ratio from our Phase 1 study provided a better fit for the Level 1 services, and that from multicentre analysis provided a better fit for the Level 2 services. These two sets of ratios were taken forward into our modelling for the weighted bed day costing model (see chapter seven).

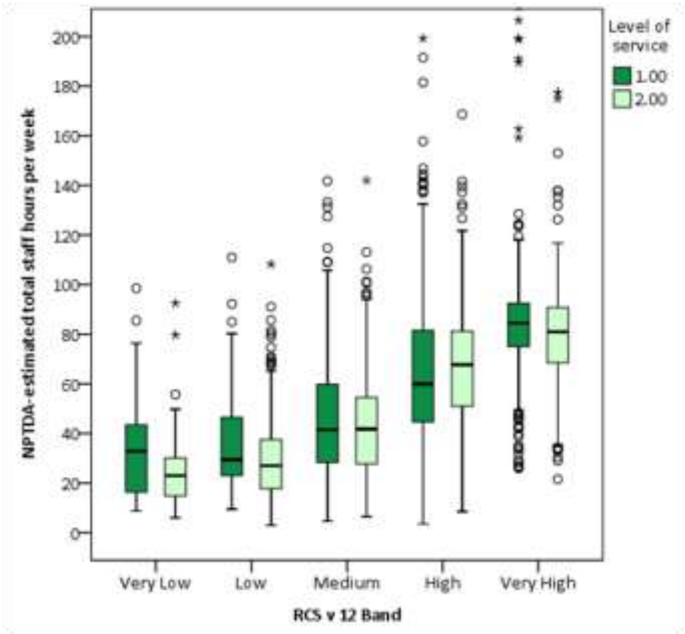



**Figure 37: The distribution of total staff hours per week for the Level 1 and 2 services**

# Analysis of caseload complexity in relation to resources (medical, nursing and therapy staff) to inform designation of specialist rehabilitation services by NHS England

In this analysis we sought to:

a) compare caseload complexity in the different categories of service

b) compare the levels of rehabilitation intervention that are typically provided for patients in different categories of complexity, across different service settings

c) define the staffing establishment levels that are currently provided to meet these different requirements.

We also explored the characteristics of rehabilitation need that define a requirement for specialist (Level1 and 2) rehabilitation services.

## *Settings*

Under the PbR guidance from 2012/13, registration with UKROC and submission of at least the minimum UKROC dataset for every case episode became a mandated commissioning requirement for all Level 1 and 2 specialist rehabilitation services.

This meant that, instead of relying on data from a pilot sample of 10-20 services, we were able to obtain data directly from all registered units, which in 2013/14 included

- 18 Level 1 (tertiary) services, commissioned by NHSE
- 16 Level 2a (supra-district) services, commissioned by NHSE
- 24 Level 2b (local district) services, commissioned by CCGs

## *Methods*

To measure caseload complexity and the staffing inputs provided to meet the needs of the caseload, we used data collated from the UKROC Service profiles (updated annually), and also the parallel ratings of the RCS-E, NPDS and NPTDA for all patients on the unit in cross-sectional tranches at approximately fortnightly intervals as follows:

- The Rehabilitation Complexity Scale (RCS-E) as a measure of complexity of rehabilitation needs.
- The Northwick Park nursing and therapy Dependency scales (NPDS/NPTDA), which translate into measures of nursing and therapy hours per week. These were summed to estimate the total staff time for the caseload during that week.

To compare these directly we used parallel ratings of the RCS-E-v12, NPDS and NPTDA and measured these three instruments in parallel in a series of cross-sectional cohorts involving all in-patients present in a given service during that week.

- Caseload complexity during each period was defined both in terms of the mean RCS-E scores and in terms of the relative proportion of the five Complexity bands (Very Low, Low, Medium, High and Very High.

- Because caseload varies over time, 4-6 cross sectional cohort analyses, a minimum of 2 weeks apart, were pooled to provide a sample size of at least 100 cases per unit providing snapshot of each service representative of a typical week.

We then compared the estimated staffing hours provided for the caseload with the total staff available (calculated from the total medical nursing and therapy staff whole time equivalent (WTE) establishment for each unit), to provide an estimate of the total staffing levels required to deliver that level of patient-related care - taking into account the proportion of time that staff spend on other (non-clinical) activities such as general administration, management, professional development, audit, appraisal etc). Services were grouped into the levels shown in *Table 68.* By the end of the programme, the increased demand for hyper-acute rehabilitation at a stage when patients still have unstable medical/surgical needs had created the need for a further Level 1 HA (Hyper-acute).

**Table 68: Service levels**

| | | |
|---|---|---|
| **Level 1 – Tertiary specialist rehabilitation services >85% category A patients** | | |
| **1a** | **Complex physical disability** | Predominantly high physical dependency including tracheostomy, ventilator patients |
| **1b** | **Mixed** | Mixed physical and cognitive disabilities |
| **1c** | **Cognitive behavioural** | Mainly 'walking wounded' patients with cognitive problems, challenging behaviour |
| **Level 2 – Specialist rehabilitation services** | | |
| **2a** | **Supra-district** | Mixed cased load with at least 50% category A patients |
| **2b** | **Local District** | Predominantly category B patients |

We originally expected to perform a more detailed analysis of the total staff time used in patient care (as opposed to other duties) from staff timetables recorded during at least two of the caseload and intervention sampling periods. The proportions of time spent in direct patient care and in other duties (admin, management education etc) were be identified and collated by the centre co-ordinator. In the event, this was not feasible given the data burden already imposed on centres by the increased reporting requirements of the new commissioning arrangements.

Detailed activity analyses previously undertaken in one unit[53, 84, 101] demonstrated an overall figure of approximately 65% of time spent in clinical care for senior therapy and nursing staff (band 7-8) with managerial responsibilities, 70-75% of time for middle grade staff (band 5-6) and 80% for care and therapy assistants (giving an overall figure of approximately 75% across all grades). A similar figure was found for the nursing staff, as described in Chapter 4 (the section on further development of the NPDS for hospital settings). These data were taken forward in our calculations.

### Relationship to policy development

Once again the analysis was conducted in several phases.

- A preliminary multi-centre analysis of the caseload data reported to UKROC up to November 2012 was performed to identify rehabilitation resources provided in relation to caseload complexity. At that point, data were available from 12 services signposted as Level 1; 9 as Level 2a and 13 as Level 2b services. A further 12 services were submitting at least some data, but for many the datasets were too small to allow accurate assessment.

  *The results were shared with NHSE and used to inform initial identification of the eligible specialist rehabilitation services for commissioning from April 2013 under the specialist rehabilitation.*

- Transition to the new commissioning arrangements and mandated registration with UKROC led to more complete data collection. Following registration of a further 22 services, multi-centre analysis of the data for the first half of 2013/14 was repeated to confirm/refute the original findings. As a result, 58 services were confirmed as Level 1 or 2 services, although some of these were re-signposted for designation at either a higher or lower level (within the range of 1a, 1b, 1c, 2a, 2b) on the basis of their service profile and activity.

  *These findings were shared with NHSE England to inform commissioning for 2014/15.*

## Final analysis

The final analysis presented here was run on the full year data for 2013/14.
*Table 69* shows the Level 1 and 2 services for adult specialist rehabilitation, together with the occupied bed day activity as reported to UKROC by the year end of 2013-14. We also analysed the estimated total annual activity reported by the units in their service profiles.

Overall, the activity reported to UKROC was 97.2% of the estimated total. The loss of approximately 3% may reflect some incomplete reporting, but also reflects the fact that UKROC collates only bed days that fall within the limit of 180 days, in accordance with the NHSE Service specification for specialist rehabilitation. (Although some regions have a system for approving extended stays beyond 180 days in exceptional cases, as yet there is no established mechanism for reporting these approvals within the de-identified UKROC dataset).

Across all the Level 1 and 2 services, the total number of occupied bed days (350K) translated into a total bed base of approximately 1117 specialist rehabilitation beds – the proportion of Level 1:2 beds was 32%:68%. However, as NHSE also commissioned all Level 2a beds in this first year of transition, approximately 58% was commissioned by NHSE England.

**Table 69: Breakdown of rehabilitation units by service level with reported occupied bed activity in 2013/14 compared with self-reported OBD per annum**

| Level of unit and UKROC ID | Service Name | No of beds | UKROC reported OBDs 2013/14 | Self-reported OBD |
|---|---|---|---|---|
| **Level 1a Hyper-acute** | | | | |
| **C130** | Walton Centre, Liverpool | **10** | 3,029 | 3,504 |
| **C182** | Salford Royal Hospital, Manchester | **12** | 4,669 | 4,205 |
| **Total 1a Hyper-acute** | | **22** | **7,698** | **7,709** |
| **Level 1a Physical Disability** | | | | |
| **C029** | Coleman Centre, Norwich, Norfolk | 20 | 6,034 | 6,643 |
| **C031** | RRU, Northwick Park Hospital, London NW | 24 | 8,403 | 8,760 |
| **C035** | Walkergate Park, Newcastle | 33 | 8,968 | 11,416 |
| **C038** | Oxford Centre for Enablement | 26 | 8,501 | 8,575 |
| **C075** | Brain Injury Unit, Leicester | 9 | 2,441 | 3,154 |
| **C090** | RHN Putney, London SW | 49 | 14,499 | 9,681 |
| **C181** | C2, Salford Royal Hospital, Manchester | 20 | 6,679 | 6,935 |
| **Total 1a** | | **181** | **55,525** | **55,164** |
| **Level 1b Mixed (physical disability and cognitive behavioural)** | | | | |
| **C016** | Frenchay, Bristol | 29 | 8,449 | 10,162 |
| **C088** | RNRU, Homerton Hospital, London NE | 27 | 7,692 | 8,004 |
| **C131** | CRU, Walton Centre, Liverpool | 20 | 5,996 | 6,074 |
| **C165** | W8, Moseley Hall Hospital, Birmingham | 8 | 2,513 | 2,686 |
| **C166** | W9, Moseley Hall Hospital, Birmingham | 24 | 8,426 | 8,060 |
| **C189** | Royal Leamington Spa Hospital, Warwick | 33 | 8,979 | 11,563 |
| **Total 1b** | | **141** | **42,055** | **46,549** |
| **Level 1c Cognitive Behavioural - no published 2013/14 tariff** | | | | |
| **C091** | Lishman Unit, London SE | 7 | 1,670 | 2,453 |
| **C137** | Walkergate Park, Newcastle | 14 | 3,160 | 3,779 |
| **Total 1c** | | **21** | **4,830** | **6,232** |
| **Level 2a Supra-district** | | | | |
| **C009** | Neuro Rehabilitation Unit, Trafford | 30 | 8,844 | 8,486 |
| **C014** | East Kent Neuro Rehabilitation Unit, Kent | 19 | 6,761 | 6,588 |
| **C022** | King's College Hospital, London | 15 | 5,167 | 5,256 |
| **C023** | Preston Neuro Rehab Unit, Lancs | 12 | 3,615 | 4,205 |
| **C040** | Dorset Brain Injury Service, Poole | 5 | 1,762 | 1,744 |
| **C041** | Phoenix Rehab Centre, Portsmouth | 12 | 4,305 | 4,114 |
| **CO59** | Sussex Rehab Centre, Sussex | 43 | 14,346 | 13,497 |
| **C064** | Haywood Hospital, Stoke on Trent | 23 | 5,394 | 7,975 |
| **CO69** | University College Hospital, London | 18 | 6,383 | 6,295 |
| **C071** | Ashby Rehab Unit, Lincs. | 12 | 4,445 | 4,205 |
| **C076** | Young Disabled Unit, Leicester | 10 | 2,875 | 2,803 |
| **C109** | Plymouth Neuro Rehab Unit, Devon | 15 | 5,361 | 5,256 |
| **C121** | Blackheath (HNDU), London SE | 17 | 4,514 | 4,748 |
| **C134** | Wolfson Neurorehab Centre, London SW | 32 | 11,543 | 9,344 |
| **C183** | Ward 7, Salford Royal Hospital | 10 | 3,597 | 3,541 |
| **Total 2a** | | **273** | **88,912** | **88,057** |

| Level of unit and UKROC ID | Service Name | No of beds | UKROC reported OBDs 2013/14 | Self-reported OBD |
|---|---|---|---|---|
| **Level 2b Local specialist rehabilitation** | | | | |
| **C002** | Addenbrooke's Hospital, Cambridge | 8 | 2,247 | 2,774 |
| **C003** | Willesden Centre, London NW | 12 | 4,322 | 4,292 |
| **C010** | Hume Neurorehab Unit, Tyne & Wear | 19 | 7,109 | 6,935 |
| **C012** | Kings Lodge, Derby | 18 | 5,624 | 6,259 |
| **C018** | Alderbourne Rehab Unit, London NW | 20 | 6,806 | 7,052 |
| **C020** | St. Mary's Hospital, Isle of Wight | 22 | 7,629 | 7,869 |
| **C021** | West Kent Neuro Rehabilitation Unit, Kent | 9 | 2,721 | 3,232 |
| **CO25** | Chapel Allerton Hospital, Leeds | 22 | 7,224 | 7,629 |
| **C026** | Pinderfields Hospital, Yorkshire | 15 | 4,300 | 4,654 |
| **C036** | Linden Lodge, Notts. | 25 | 7,958 | 7,621 |
| **C043** | Oakwood, Rotherham | 9 | 1,733 | 2,628 |
| **C044** | Royal Berkshire Hospital, Reading | 16 | 5,580 | 5,840 |
| **C049** | Donald Wilson House, Chichester, W. Sussex | 11 | 3,627 | 3,922 |
| **C053** | Osborn Unit, Sheffield | 14 | 4,902 | 4,490 |
| **C057** | Snowdon Unit, Southampton | 15 | 4,830 | 4,458 |
| **C063** | Devonshire Unit, Stockport | 19 | 5,917 | 6,368 |
| **CO65** | Bradley Unit, Woking | 17 | 6,451 | 6,132 |
| **C067** | Taunton Neuro-Rehab Service, Somerset | 12 | 3,857 | 4,205 |
| **CO68** | Floyd Unit, Rochdale | 18 | 5,913 | 5,621 |
| **C073** | Neurorehab Unit, Coventry | 12 | 4,126 | 4,122 |
| **C080** | Clatterbridge Rehabilitation Centre, Wirral | 7 | 1,916 | 2,173 |
| **CO81** | West Park, Wolverhampton | 10 | 3,644 | 3,342 |
| **C082** | Leigh Infirmary, Wigan & Leigh | 22 | 6,065 | 6,456 |
| **CO92** | Airedale Neurorehab Services, Yorkshire | 10 | 1,500 | 1,825 |
| **C095** | Royal Free Neurorehab Centre, London NC | 11 | 3,442 | 3,199 |
| **CO98** | W 26, James Cook University Hospital, S Tees | 16 | 5,160 | 5,144 |
| **C101** | Albany Rehab Unit, London NC | 10 | 3,329 | 3,266 |
| **C105** | Cumberland Infirmary, Cumbria | 10 | 2,087 | 2,555 |
| **C120** | Mardon Neuro-Rehab Centre, Exeter | 11 | 4,136 | 3,854 |
| **C132** | Broadgreen Hospital - Liverpool Network | 15 | 3,921 | 4,106 |
| **C133** | St Helens Hospital, Merseyside | 20 | 6,244 | 7,008 |
| **C015** | Rakehead Rehab Centre, East Lancashire | 17 | 4,358 | 5,500 |
| **C135** | Queen Mary's Hospital, Roehampton | 14 | 4,046 | 4,344 |
| **Total 2b** | | **486** | **152,724** | **158,875** |
| **Grand Total** | | **1,124** | **351,744** | **362,586** |

*Table 70* shows the mean % of serial RCS-E ratings in each band, within the different service levels. *Table 71* shows the mean nursing and therapy dependency scores and the estimated staff hours per week for each occupied bed. *Table 72* shows the mean staffing levels (WTE per occupied bed) within the different services. We expected and found a progressive increase in the complexity profile and nursing and therapy dependency across the levels from 2b to 1a, with a corresponding increase in staffing levels to meet the needs of the caseload.

**Table 70: Mean % of serial RCS-E ratings in each band, by level of service**

| | | | Complexity Band | | | |
|---|---|---|---|---|---|---|
| **Level** | **Very Low %** | **Low %** | **Medium %** | **High %** | **Very High %** | **% RCS-E v12 ≥11** |
| **1a** | 3.9% | 2.2% | 14.2% | 52.3% | 27.4% | 79.7% |
| **1b** | 2.0% | 2.2% | 23.4% | 50.6% | 21.8% | 72.4% |
| **1c** | 8.8% | 0.3% | 25.3% | 43.7% | 21.9% | 65.6% |
| **2a** | 3.9% | 4.0% | 26.0% | 51.8% | 14.4% | 66.1% |
| **2b** | 9.8% | 6.9% | 28.1% | 45.9% | 9.3% | 55.2% |

**Table 71: Mean nursing (NPDS) and therapy dependency (NPTDA) scores and the estimated staff hours per week for each occupied bed by level of service**

| Level 1 | Mean NPDS total score | Mean NPDS-estimated total care/nursing hours | Mean NPTDA total score | Mean NPTDA-estimated total therapy hours | Mean total staff hours |
|---|---|---|---|---|---|
| **1a** | 47.1 | 24 | 36.5 | 29 | 53 |
| **1b** | 40.7 | 22 | 21.2 | 25 | 47 |
| **1c** | 23.1 | 14 | 45.3 | 37 | 51 |
| **2a** | 38.6 | 20 | 21.2 | 21 | 41 |
| **2b** | 37.1 | 19 | 19.9 | 19 | 38 |

**Table 72: Mean staffing levels (WTE per occupied bed) by level of service**

| Level | Nursing staff WTEs/Bed | Therapy staff WTEs/Bed | Medical staff WTEs/Bed | Other staff WTEs/Bed | Total staff WTEs/Bed |
|---|---|---|---|---|---|
| **1a** | 2.1 | 1.1 | 0.2 | 0.3 | 3.7 |
| **1b** | 2.0 | 1.1 | 0.2 | 0.4 | 3.6 |
| **1c** | 3.3 | 0.7 | 0.3 | 0.3 | 4.6 |
| **2a** | 1.8 | 0.9 | 0.2 | 0.2 | 3.1 |
| **2b** | 1.7 | 0.9 | 0.2 | 0.2 | 3.1 |

*Table 73* assimilates the above information into a set of standard profiles for each level of service. In 2014, UKROC supplied each provider of a Level 1 or 2 specialist rehabilitation with a report of their unit's profile in comparison with the standard for their group. Figure 38 shows an example of such a report.

Where outliers were identified (i.e. substantial variance form the norm), further discussion to place with the provider to try to determine whether this was a genuine difference or a matter of over-under-scoring of one or more of the instruments. In cases of mis-scoring further training as offered, followed by review of activity reported over the ensuing 3-6 months to determine whether the unit had moved closer to the norm.

**Table 73: Standard service profiles for each level of service**

| Mean | 1a n=7 | 1b n=8 | 2a n=14 | 2b n=23 |
|---|---|---|---|---|
| | Mean | Mean | Mean | Mean |
| **Staffing WTE/OB's pa** | | | | |
| Therapy | 1.1 | 1.0 | 0.8 | 0.8 |
| Nursing/care | 2.1 | 2.0 | 1.9 | 1.8 |
| Medical | 0.2 | 0.2 | 0.2 | 0.2 |
| **Total** | **3.4** | **3.2** | **2.9** | **2.8** |
| Weighted Cost per OBD | £411 | £382 | £358 | £366 |
| Cost per OBD | £540 | £483 | £452 | £418 |
| **Dependency (hours) -E** | | | | |
| NPTDA | 29.8 | 22.6 | 21.6 | 19.7 |
| NPDS | 44.8 | 40.4 | 39.8 | 39.4 |
| **Total** | **74.6** | **63.0** | **61.4** | **59.1** |
| **Complexity** | | | | |
| RCS-Ev12 | 12.7 | 12.1 | 11.5 | 10.5 |
| **%RCS: 11-20** | **81.5%** | **71.3%** | **66.2%** | **53.4%** |

**Comparisons Table For:** Cxx
**Current Level:** 2a
**Indicative Level:** 2a

**Your Service Data**
*(below OBD's & Costs are based on full year 13/14, month 6 out-turn)*

| | |
|---|---|
| No. of Beds at start of year (information from Provider): | 32.0 |
| No. of Beds at end of year (information from Provider): | 32.0 |
| Average Bed Base pa: | 32.0 |

| WTEs | | | |
|---|---|---|---|
| | Calculated OB's / pa | 32.1 | |
| 29.3 | Reported OBDs | 11,732 | ie 100.4% occupancy |
| 36.9 | Weighted OBDs | 16,112 | |
| 3.9 | Cost excl MFF | £3,939,923 | |
| 70.1 | MFF% | 21.25% | |

| Mean | Comparater Group 2a Mean | Your Service | |
|---|---|---|---|
| **Staffing WTE/OB's pa** | | | |
| Therapy | 0.8 | 0.9 | |
| Nursing/care | 1.9 | 1.1 | Very Low |
| Medical | 0.2 | 0.1 | Low |
| Total | 2.9 | 2.2 | Very Low |
| Weighted Cost per OBD | £358 | £245 | Very Low - but reflects low staffing levels |
| Cost per OBD | £452 | £336 | Mean Cost per OBD based on the number of OBD's as reported through UKROC (ie excluding >180 days) |
| | | Very Low | |
| **Dependency (hours) -E** | | | |
| NPTDA | 21.6 | 26.5 | Higher than average - but reflects high therapy staff levels |
| NPDS | 39.8 | 29.5 | Very Low - but reflects low staffing levels |
| Total | 61.4 | 56.0 | Very Low |
| **Complexity** | | | |
| RCS-Ev12 | 11.5 | 12.5 | Very high - likely to reflect over-reporting given low staffing levels |
| %RCS: 11-20 | 66.2% | 87.1% | Very High |
| | | 88.4% | RCS 11-20% excluding the unbanded bed days effect. Very High |
| | | 0.0% | % of Total OBDs that are > 180 |

**Therapy Hrs**

| | |
|---|---|
| 851.8 | Per week total Therapy hrs (based on reported NPTDA hrs) |
| 29.1 | ie at this rate, each therapist is being recorded as putting in 29.1hrs / week actual therapy time with patients |
| 77.5% | If this NPTDA Mean of 26.5 is a true reflection of therapy hrs, **77.5%** of each therapist's time would be patient contact time |

**Nursing Hrs**

| | |
|---|---|
| 948.2 | Per week total Nursing hrs (based on reported NPDS hrs) |
| 25.7 | ie at this rate, each Nurse is being recorded as putting in 25.7hrs / week actual therapy time with patients |
| 68.5% | If this NPDS Mean of 29.5 is a true reflection of nursing hrs, **68.5%** of each Nurse's time would be patient contact time |

**Figure 38: Example of feedback to a service provider on their service profile data**

# A system for identifying patients with category A needs (requiring specialist (Level 1/2a) services

The NHSE service specification requires that at least 85% of patients in a Level 1 service have category A needs, whilst a Level 2a service was expected to have at least 50% category A patients. Going forward, only category A patients will continue to be commissioned by NHS England. Therefore it was necessary to identify the characteristics of rehabilitation need that define a requirement for Level 1 and 2 services.

The primary definitions of category A and B need for Level 1 and 2 services were laid down in the Department of Health's SSNDS definition set for Brain injury and complex rehabilitation,[75] and subsequently carried through to the NHSE service specification.[44] These are more specific than the general resource requirements that are measured by the RCS-E. Therefore a checklist was drawn up collect these data, and subsequently a 3-level scoring system was added to form the Patient Categorisation (PCAT) tool. The itemised tool is operational within the UKROC database update released in April 2014, and from 2014/15, all Level 1/2a specialist rehabilitation providers commissioned by NHSE are required to record this to confirm that patients admitted to these services have category A or B needs.

## *PCAT scores within the different levels of service*

The purpose of the PCAT Tool is to distinguish between patients with different levels of need and so direct them to the appropriate level of service. Whilst we would anticipate that all services have a certain mix of category A and B patients, we would expect the PCAT tool to reflect a higher proportion of Category A patients in the Level 1 services compared with Level 2.

In chapter 4 we described an evaluation of the psychometric properties of the PCAT tool. The findings suggested that a PCAT total score ≥30 was a reasonable indicator of category A needs. In the following analysis we examine PCAT scores across different levels of service in order to:

a) Describe the range of PCAT scores within the different service levels

b) Use the PCAT tool to identify the proportion of patients with category A needs in each service level, as defined by a PCAT total score ≥30

The following analysis was conducted on all PCAT data recorded in the UKROC database for all Level 1 and 2 services between April 2012 and December 2014. During this period, PCAT ratings were recorded on admission for a total of 5396 patients, comprising 63.1% of the cohort. The demographics of this sample were: mean age 54.4 (sd 18.2) years Males: females 58:42%; Diagnostic distribution – acquired brain injury 66.4%, spinal cord injury 9.4%, peripheral neurology 5.0% and progressive conditions 9.9%.

*Table 74* and *Figure 39* show the comparative distribution figures across the different levels of service.

**Table 74: Mean, median and range for total PCAT scores on admission across the different levels of service**

| | N | Mean | Std. Dev | Median | IQR | Range |
|---|---|---|---|---|---|---|
| **Hyper-acute** | 81 | 38.5 | 4.8 | 39 | 35-43 | 24-47 |
| **1a** | 578 | 33.5 | 6.2 | 34 | 29-38 | 16-48 |
| **1b** | 829 | 31.6 | 7.8 | 31 | 25-38 | 17-48 |
| **1c** | 61 | 35.7 | 6 | 36 | 32-40 | 18-47 |
| **Paediatric** | 79 | 38.8 | 5.3 | 40 | 46-42 | 17-47 |
| **2a** | 1410 | 28.4 | 5.8 | 28 | 24-32 | 16-46 |
| **2b** | 2358 | 28.2 | 6.3 | 28 | 23-33 | 16-48 |

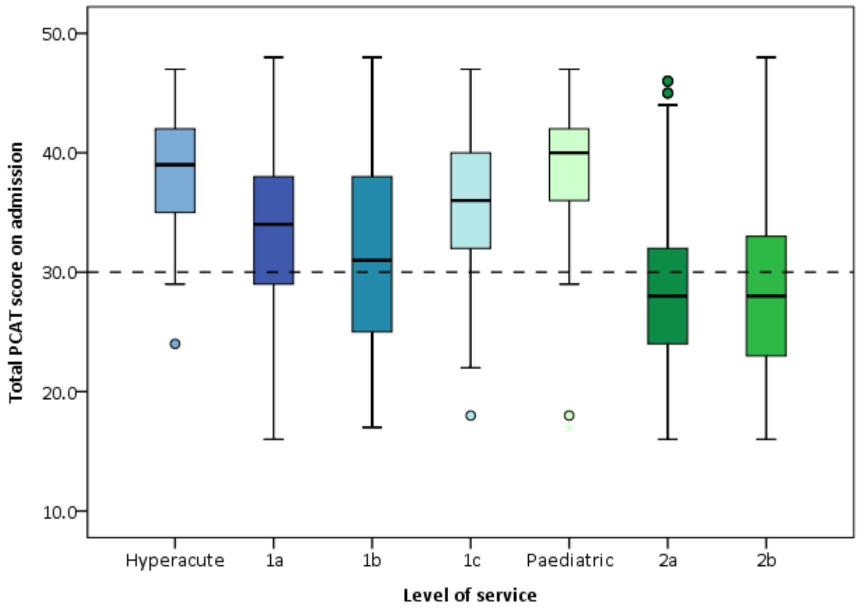



**Figure 39: Frequency distribution of total PCAT scores across the different service levels**

As expected, the Level 1 services had higher mean PCAT scores than the Level 2 services, and the figures also suggest that a mean PCAT score of 30 or more may provide a reasonable indicator of services that have a high proportion of patients with category A needs.

*Figure 40* shows the distribution of PCAT sub-scale scores across the different service levels. The results mirror the pattern of PCAT Total scores quite closely except for the Level 1c (cognitive-behavioural) services, which have substantially higher PCAT-Cog scores, as would be expected.

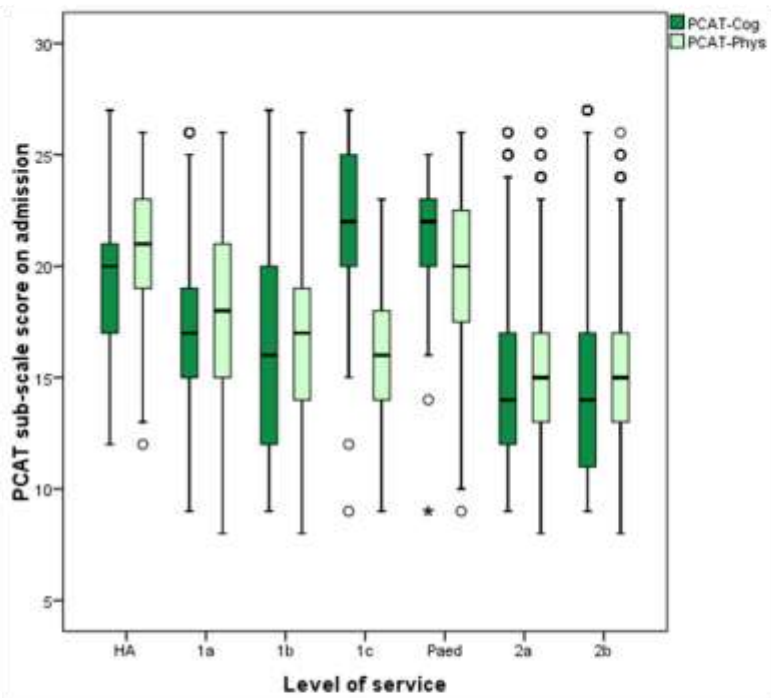



**Figure 40: The distribution of PCAT subscale scores across the different service levels**

*Figure 41* shows the percentage of category A patients across the different levels of service a) as defined by a PCAT score ≥30 and b) as defined but assessor subjective ratings. The relative proportions were very similar although there was a tendency for subjective ratings to identify a higher proportion of patients as having category A needs in the Level 1b services and a lower proportion in the Level 2 services. Although we expected to find the PCAT a less sensitive tool for the Level 1c (cognitive behavioural) services, it actually performed quite well.

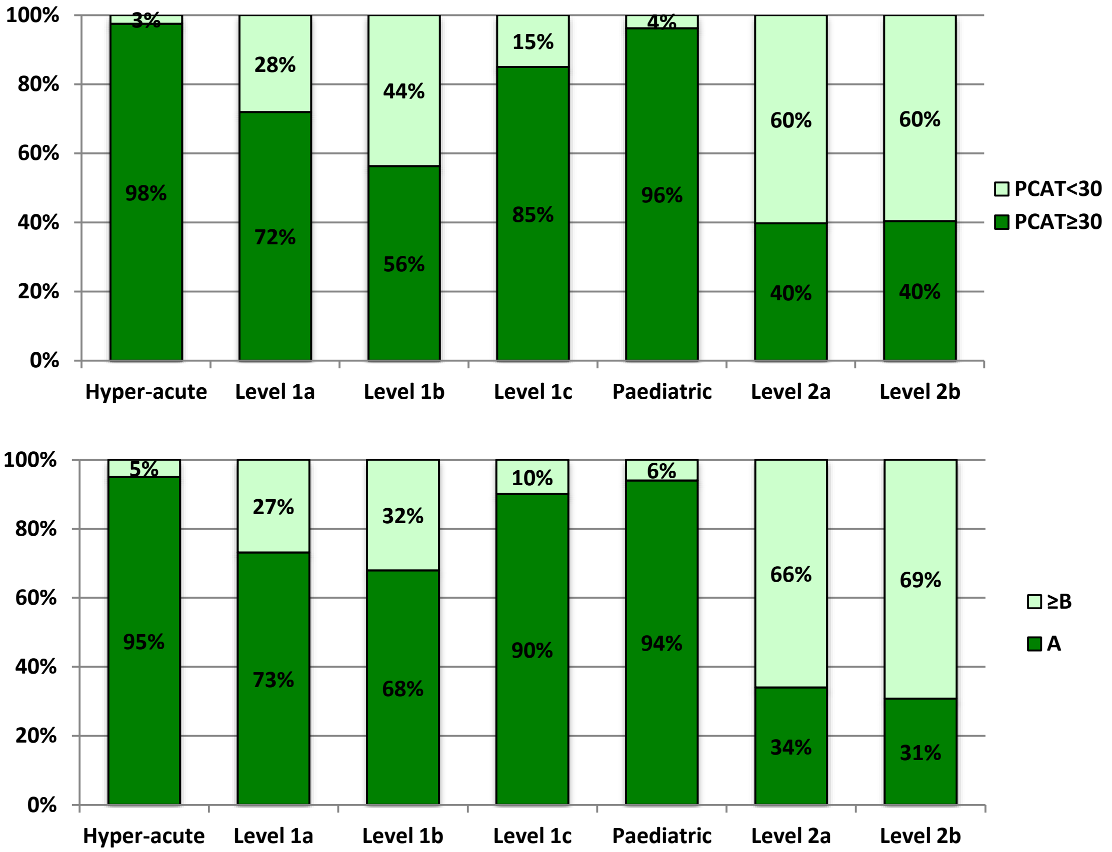



**Figure 41: Percentage of category A patients, across the different levels of service a) as defined by a PCAT score ≥30**

## Discussion

Taken together, these findings confirm that, as expected, the Level 1 services have a significantly higher proportion of patients with category A needs, whether defined by the PCAT tool or by subjective assessment. The data provide supporting evidence for the PCAT as a useful tool for distinguishing the complexity of patients needs in the different levels of service, and a cut off point of PCAT Total score ≥30 appears to provide a reasonable working threshold for category A needs.

The data suggest that some Level 1a, 1b and 2a services are currently taking a higher proportion of category B patients than the specification demands. This resonates with clinical experience, as it is recognised that many Level 1/2a services are currently commissioned at prices well below cost, and are thus constrained to admit a proportion of category B patients as they are not sufficiently resourced to manage the demands of a more complex caseload.

# Medical needs and hyper-acute care - Further analysis of MAA data

The growing body of evidence for the effectiveness of early intensive rehabilitation has led to increasing recognition of the need for rehabilitation at start at an early stage in the recovery pathway. Introduction of the Rehabilitation Prescription within the major trauma pathway, has led to more active involvement of rehabilitation medicine consultants within the major trauma centres, which is now expanding to other areas of care including acute stroke and neurosciences.

Pressure to move patients into specialist rehabilitation services as soon as their immediate care and treatment is complete has led to the development of 'hyper-acute' rehabilitation services – that is services that can take patients at a stage in their recovery when they may still have unstable medical / surgical needs. Even for those who have stabilised, there may still be a need to complete investigations and procedures that would previously have been undertaken in the acute services prior to transfer.

The development of these services relies on accurate information about the additional resources required to manage the medical needs of these patients safely and effectively.

The Rehabilitation Complexity Scale M score and associated Medical Activities Assessment have been developed to capture information on resource use including:

- The frequency and nature of medical interventions, including investigations and procedures
- The involvement of other specialities to inform co-dependencies
- The medical resources required including medical time during working hours as well as out-of-hours emergency medical/surgical care or cover.

The Regional Rehabilitation Unit at Northwick Park became formally designated as a Level 1a/hyper-acute rehabilitation service from April 2014, but at the outset it was uncertain exactly how what proportion of its beds should be commissioned on the hyper-acute side.

Following on from the initial piloting described in Chapter 4, serial data collection continued from March 2014 to March 2015, applying the MAA on a weekly basis for all in-patients admitted to this service. The objectives of the analysis presented below were to describe and quantify the range of medical activity within a Level 1a/ hyper-acute rehabilitation service, and so to define the principal co-dependencies on other services / specialties and the additional requirements for medical time and emergency cover.

### Study sample

A total of 1209 data points were captured during the periods from a total of 101 patient episodes. The demographics of the study sample were as follows:

- Mean age 43.5 (SD13) range 17-66 years. Males: females 65:35%
- Mean Length of stay was 106 days (sd 51) range 6-337 days.
- Diagnosis: Acquired brain injury 94 (93%), Spinal Cord Injury 3 (3%), Peripheral Neurology 3 (3%); other 1(1%)

### Data capture:

- RCS-E v 13 scores and the MAA tool were completed for each patient every week and verified by consensus of the RRU medical team during the weekly medical ward round in order to maximise accuracy.
- Medical time from the RRU team was recorded through the NPTDA medical scores.

### Results

As expected there was a reasonably strong correlation (Spearman rho 0.48, p<0.001) between the RCS-M score and hours of medical intervention. The frequency of RCS-M scores and the mean hours of medical intervention by the RRU medical team are shown in *Table 75* and *Figure 42*. Patients with an RCS-M score of 3 consumed an average of 2-3 hours of medical time per week, and those with M scores of 4 consumed an average of 3-4 hours, but with a range of up to 6 hours.

**Table 75: RCS M score frequency and associated medical time**

| RCS-M score | Frequency N (%) | Medical hours Mean (SD) |
|---|---|---|
| M0 | 0 | - |
| M1 | 185 (15.3%) | 1.5 (0.6) |
| M2 | 544 (45.0%) | 2.1 (0.8) |
| M3 | 347 (28.7%) | 2.6 (1.1) |
| M4 | 133 (11.0%) | 3.4 (1.4) |
| | 1209 | |

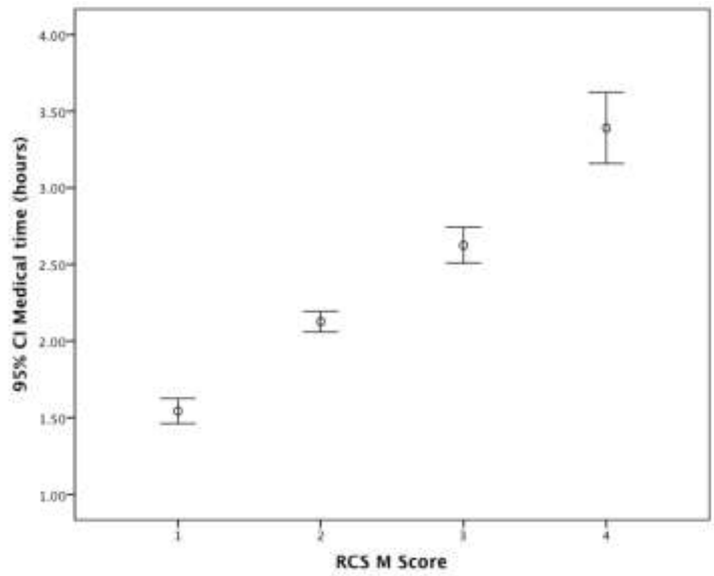



**Figure 42: Mean medical time with 95% confidence intervals in relation to RCS-M score**

### Nature of interventions.

As in the pilot study the range of interventions was diverse, and spanned a number of medical activities that had not been captured in the pilot study. This evaluation has been used to expand further the options within the MAA tool.

The frequency of routine medical issues addressed is shown in *Table 76*. Routine review and medication management was undertaken for most patients every week. Other commonly addressed issues were pain management, medical management of bladder and bowel incontinence and addressing abnormal blood tests.

**Table 76: Frequency of routine medical issues addressed**

| Routine clinical issues | N | (%) |
|---|---|---|
| Routine Review | 1097 | (91%) |
| Medication Management | 921 | (76%) |
| Pain Management | 271 | (22%) |
| Bowel Management | 238 | (20%) |
| Bladder Management | 138 | (11%) |
| Abnormal blood tests | 119 | (10%) |
| New Admission | 87 | (7%) |
| Tone issues | 83 | (7%) |
| Possible/confirmed sepsis | 97 | (8%) |
| Behavioural Management | 83 | (7%) |
| Complex nutritional Management | 57 | (5%) |
| Warfarin Management | 53 | (4%) |
| Mood Management | 52 | (4%) |
| Diabetes Management | 43 | (3%) |
| Venous thrombo-embolism | 24 | (2%) |
| Psychological/psychiatric risk assessment | 18 | (2%) |

The frequency of interventions and procedures is shown in *Table 77*

- 850 data-points (70% of the total) included at least one investigation, of which the commonest, unsurprisingly, were blood tests, microbiology and imaging.
- 197 (16%) of data-points included at least one procedure, of which the commonest were intravenous cannulation, and then various interventions for tracheostomy, enteral feeding and spasticity management.
- Other important procedures that occurred less commonly were neurosurgical, orthopaedic, urological, ophthalmological; and ENT procedures.

Despite our efforts to capture the information systematically, some notable interventions were missing from the list - including out of hours replacement/checking of nasogastric tubes whilst waiting for gastrostomy. These are known to be a frequent occurrence on this unit – so frequent in fact that they fell beneath the threshold for medical vigilance this series.

**Table 77: Frequency of investigations and procedures**

| Investigations | N | % |
|---|---|---|
| Bloods | 429 | 50.5% |
| Microbiology | 189 | 22.2% |
| X-Ray | 110 | 12.9% |
| CT | 42 | 4.9% |
| Ultrasound | 28 | 3.3% |
| ECG | 13 | 1.5% |
| MRI | 9 | 1.1% |
| Endoscopy | 8 | 0.9% |
| Urology | 7 | 0.8% |
| Standard Echo | 6 | 0.7% |
| Lung Function | 3 | 0.4% |
| EEG | 2 | 0.2% |
| 24 hour ECG | 2 | 0.2% |
| Bubble Echo | 1 | 0.1% |
| EMG | 1 | 0.1% |
| **Total** | **850** | **70%** |

| Procedures | N | % |
|---|---|---|
| Intravenous cannulation | 83 | 42.1% |
| Tracheostomy removal | 21 | 10.7% |
| Gastrostomy insertion | 15 | 7.6% |
| Toxin injection | 13 | 6.6% |
| Gastrostomy removal | 11 | 5.6% |
| Tracheostomy weaning | 9 | 4.6% |
| Tracheostomy change | 5 | 2.5% |
| Shoulder injection | 4 | 2.0% |
| Difficult urethral catheterisation | 4 | 2.0% |
| VP shunt | 4 | 2.0% |
| Orthopaedic surgery/tenotomy | 4 | 2.0% |
| Chest drain/pleural tap | 3 | 1.5% |
| Ophthalmic procedure | 3 | 1.5% |
| S/P catheter change/insertion | 3 | 1.5% |
| ENT procedure | 2 | 1.0% |
| IV filter removal | 2 | 1.0% |
| Other joint injection | 2 | 1.0% |
| Surgical wound debridement | 2 | 1.0% |
| Caesarian Section | 1 | 0.5% |
| CPAP | 1 | 0.5% |
| Cranioplasty | 1 | 0.5% |
| Laparotomy | 1 | 0.5% |
| Laser treatment for granulation | 1 | 0.5% |
| Endoscopic treatment | 1 | 0.5% |
| Reduction of paraphimosis | 1 | 0.5% |
| **Total** | **197** | **16%** |

*Table 78* shows the frequency of input from other specialties.

- The most common requirements were for input from ENT ( including tracheostomy support) neurology, neurosurgery, obstetrics, orthopaedics, gastroenterology, ophthalmology, urology and acute medicine.
- Other commonly required inputs were from cardiology, endocrinology, respiratory medicine, palliative care, rheumatology, microbiology and neurophysiology.
- The list of less commonly involved specialties is also important, as these patients do not travel easily, so unless these supports are readily available to come to the patient’s bedside, they are likely to miss out on important aspects of treatment.
- Support from ITU for emergency care, periarrest and ventilator support is increasingly important and a review of the medical records suggested that these may have been significantly under-reported in this dataset.

**Table 78: Requirement for input from other specialties**

| Specialty | N | % |
|---|---|---|
| **Most common** | | |
| Tracheostomy team | 126 | 22.6% |
| Neurology | 72 | 12.9% |
| Neurosurgery | 39 | 7.0% |
| Radiology | 39 | 7.0% |
| ENT | 34 | 6.1% |
| Obstetrics | 29 | 5.2% |
| Orthopaedic | 25 | 4.5% |
| Gastroenterology | 21 | 3.8% |
| Ophthalmology | 18 | 3.2% |
| Urology Team | 17 | 3.1% |
| Acute Medicine | 16 | 2.9% |
| General surgery | 9 | 1.6% |
| Interventional Radiology | 9 | 1.6% |
| Endocrinology | 8 | 1.4% |
| Cardiology | 8 | 1.4% |
| Respiratory | 7 | 1.3% |
| Oncology/Palliative care | 7 | 1.3% |
| Rheumatologist | 7 | 1.3% |
| Microbiologist | 6 | 1.1% |
| Clinical Neurophysiology | 6 | 1.1% |

| Specialty | N | % |
|---|---|---|
| **Less common** | | |
| Haematology | 5 | 0.9% |
| ITU/ventilatory | 5 | 0.9% |
| Plastics Tissue Viability | 5 | 0.9% |
| Diabetology | 5 | 0.9% |
| Gynaecology | 4 | 0.7% |
| Renal Team | 4 | 0.7% |
| Psychiatric input | 4 | 0.7% |
| Acute stroke team | 4 | 0.7% |
| Max –fax team | 4 | 0.7% |
| Anaesthetic Team | 3 | 0.5% |
| Audiology | 3 | 0.5% |
| Clinical Pharmacology | 2 | 0.4% |
| Infectious diseases | 2 | 0.4% |
| Dermatology | 1 | 0.2% |
| Vascular Team | 1 | 0.2% |
| A&E team | 1 | 0.2% |
| Neonatal team | 1 | 0.2% |
| **Total** | **557** | **46%** |

## *Hyperacuity*

*Table 79* records the proportion of patients who were considered (either 'definitely' or 'possibly') to require hyper-acute rehabilitation.

- 'Definite' hyper-acute needs were defined as an actual requirement for out of hours emergency medical / surgical intervention.
- 'Possible' hyper-acute needs refers to the patient being in an unstable condition, so that from a clinical perspective, safe management required them to be in a hospital environment in which emergency medical/surgical care was immediately available, even though they did not actually require such intervention during the given period.

Out of 26 beds, a total of 236 (27.4% - equivalent of 6-7 beds) had either 'definite' or 'possible' requirements for hyper-acute care. This compares with the 4/24 beds that were commissioned as hyper-acute beds by NHSE during this period. The hyper-acute patients required approximately twice the allocation of routine direct medical care, as well as on-call emergency medical and surgical support out-of-hours.

**Table 79: Frequency of hyper-acute needs and the associated medical needs**

| Hyper-acute | Frequency | | Medical hours | | RCS-M Score | |
|---|---|---|---|---|---|---|
| | N | (%) | Mean | (SD) | Median | (IQR) |
| Yes | 175 | (20.3%) | 3.0 | (1.2) | 4 | (3-4) |
| Possibly | 61 | (7.1%) | 2.6 | (0.9) | 3 | (3-3) |
| No | 624 | (72.6%) | 1.9 | (0.7) | 2 | (2-3) |

*Table 80* shows a cross-tabulation of RCS M scores with hyper-acute needs, confirming a relationship between the RCS-M score and hyper-acuity. An RCS-M score of 3-4 predicted hyper-acute care needs with sensitivity of 97% and a specificity of 74%. The positive predictive value was 58% and the negative predictive value 99%

The on-call teams were alerted to patients' medical condition on 41 occasions; 34 of which also required out of hours intervention. The out of hours team were required on an additional 36 occasions when alerting the on-call team had not occurred. Thus the RCS-M score is not by itself a very reliable indicator of actual instability and this highlights the need for some flexibility to be able to manage patients with potentially unstable needs in an environment where they still have access to medical /surgical support if needed, even though it is not an immediate requirement in that particular period.

**Table 80: Cross-tabulation for RCSM score and Hyper-acute needs**

| | Hyper-Acute? | | | |
|---|---|---|---|---|
| **RCS-M** | **Yes** | **Possibly** | **No** | **Total** |
| **4** | 94 | 8 | 5 | 107 |
| **3** | 76 | 52 | 160 | 288 |
| **2** | 5 | 1 | 318 | 324 |
| **1** | 0 | 0 | 141 | 141 |
| **Total** | 175 | 61 | 624 | 860 |

| | Hyper-acute? | | | |
|---|---|---|---|---|
| **RCS-M** | **Potential** | **No** | **Total** | |
| **3-4** | 230 | 165 | 395 | **58%** |
| **1-2** | 6 | 459 | 465 | **99%** |
| | 236 | 624 | 860 | |
| | **97%** | **74%** | | |

*Chi square 532.9 df=6 P<0.001*

## Frequency of Potential Hyper-acute Clinical issues

A total of 188 data points indicated a hyper-acute medical problem. *Table 81* shows the type and frequency of hyper-acute clinical problems for which the on-call teams were required. More than one clinical issue may have been selected but complex tracheostomy management, sympathetic storming, desaturation, respiratory/acute distress and a decline in consciousness level were the most frequently selected.

**Table 81: The frequency of hyper-acute medical problems for which intervention was required**

| Potential Hyper-acute Medical issue | Number of occasions (% of total data points n=188) | |
|---|---|---|
| **Complex tracheostomy management** | 143 | (76%) |
| **Desaturation** | 30 | (16%) |
| **Acute distress** | 30 | (16%) |
| **Respiratory Distress** | 29 | (15%) |
| **Sympathetic storming** | 26 | (14%) |
| **Decline in consciousness level** | 25 | (14%) |
| **Acute Sepsis** | 22 | (12%) |
| **Assisted Ventilation** | 19 | (10%) |
| **Seizure** | 14 | (7%) |
| **Pulmonary Embolus** | 13 | (15%) |
| **Autonomic dysreflexia** | 6 | (3%) |
| **Simple tracheostomy management** | 6 | (3%) |
| **Unstable blood sugars** | 5 | (3%) |
| **Unstable blood pressure** | 4 | |
| **Cardiac Arrest Call** | 2 | |
| **GI Bleed** | 1 | |

**Additional potential hyper-acute clinical issues included:**

- Abdominal pain x 4
- Further stroke/bleed x 5
- Pregnancy issues x 4
- Shunt issues x 5

**In conclusion:** The need for dedicated hyper-acute rehabilitation beds is increasingly recognised, and the RCS-M score and MAA provide a useful source of information to identify the patients who require these beds, and to quantify and categorise the facilities that need to be available.

However, individual patients may continue to fluctuate between medical stability/instability in the days and weeks following severe injury, and the need for hyper-acute care tends to fluctuate, so the actual number of beds required may vary over time. A flexible arrangement whereby hyper-acute beds are provided as needed within the context of a Level 1 unit in an acute hospital setting is appropriate – this is the model that operates on the Regional/Hyper-acute Rehabilitation Unit at Northwick Park.

The evidence presented here suggests that hyper-acute needs should be staffed with approximately twice the allocation of routine direct medical care, as well as the consistent availability of round the clock emergency medical and surgical care.

## *In summary for this Chapter*

Within this chapter, we set out to apply the tools in a variety of neurorehabilitation settings and to identify the rehabilitation resources currently provided in relation to caseload complexity.

We have achieved the three key deliverables which were as follows:

1. Analysis of staff time in relation to complexity of rehabilitation needs - this information was used (as described in Chapter 7) to inform development of the weighted payment model.

2. Analysis of caseload complexity in relation to resources (medical, nursing and therapy staff) – this information has been used by NHS England to inform the designation of specialist rehabilitation services.

3. A system for identifying patients with category A needs (requiring specialist (Level1/2a) services. The PCAT tool has proved an appropriate tool to quantify and describe Category A needs for rehabilitation. The RCS-M score and its associated Medical Activities Assessment has been used to characterise the requisite features of a Hyper-acute rehabilitation service.

## *Changes from the original application.*

1. We expected to work with 10 pilot centres, but due to earlier than expected national roll-out of the programme, parallel cross-sectional data has been collected from over 60 providers across a range of service configurations. The extended dataset has improved the quality of the dataset enabling sub-classification of services to give more precise specification of services and their respective costs.

2. We originally expected to perform more detailed analysis of total staff time available for patient care, as opposed to other duties, from staff timetables recorded during at least two of the caseload and intervention sampling periods. The proportions of time spent in direct patient care and in other duties (admin, management education etc) were be identified and collated by the centre co-ordinator. In the event, this was not feasible given the data burden already imposed on centres by the increased reporting requirements of the new commissioning arrangements.

On the other hand, what we had instead was systematically-reported NPDS/NPTDA data from all the level 1 and 2 centres, which proved feasible to collect within routine clinical practice. The estimations of staff time derived from the computerised algorithms within these tools resonated with clinical experience. Whilst they may not offer a completely accurate assessment of staff time in absolute terms, they provide an acceptable indication of the relative proportions of staff time across different groupings (such as service levels or complexity bands). These relative values are sufficient for the intended purposes within this programme.

3. We originally intended to use questionnaire surveys to gather the perceptions of clinical teams regarding the adequacy of staffing levels, service provision etc. Instead, this information was gathered through interactive dialogue through the various road shows. This approach proved more useful as it was possible to gain a more detailed understanding of where the shortfalls lay, together with the reasons for them.

4. We expected to be able to provide data on staffing, complexity and costs in community-based residential rehabilitation services - including specialist nursing homes and slow-stream rehabilitation units. In the absence of central commissioning with its drive for national benchmarking of these services, data collection has not been mandated in the way that it has for level 1 and 2 services. Therefore, although we have some data that will be presented in Chapter 7, it is not sufficiently complete to present a robust comparison with the Level 1 and 2 services. Nevertheless, the information gathered through this programme has contributed to the development of the BSRM Guidance to best practice for Specialist Nursing Home care for patients with complex disabilities[74].

# Chapter 6: Learning from international costing and case-mix methods

## Part 3 Objective:

To compare different international funding models and patient-level costing and case-mix methods for rehabilitation.

## Key deliverables

1. A review of international systems for case-mix and payment models in rehabilitation.
2. Comparative analysis of UK and Australian datasets across a range of long-term neurological conditions - case-mix (according to the Australian (AN-SNAP) classification), length of stay, outcomes (FIM) and costs will be complete based on the 2013/14 UKROC data.
3. Development of the national training and accreditation programme for the UK FIM+FAM and other tools in the UKROC dataset to ensure their consistent application.

In this part of the programme we examined international approaches to case-mix and payment models that are already successfully running in other parts of the world, in order to assess their potential for applicability in the UK, and to learn from their advantages and limitations.

This section describes work that has been published in two papers.[131, 186]

### *International case-mix models*

At the outset of the programme we conducted an exploratory narrative review of current case-mix and payment models in rehabilitation which was published in Clinical Rehabilitation[186]. The summary below includes an abbreviated extract from that article.

The purpose of the review was:

a) to provide an overview of the development of case-mix in rehabilitation,
b) to describe the key characteristics of some well-established case-mix and payment models in operation around the world and
c) to explore the lessons that may be learned from them and opportunities for future development towards improved models.

#### Methods

We conducted a broad-based search of the major databases (MEDLINE 1950 - November 2010 and EMBASE 1980- November 2010) using ‘rehabilitation’ in combination with ‘case-mix’ and ‘payment systems’ and related terms.  We also searched the internet and used our combined knowledge of the various systems and key authors to expand the reference base.

### Development of case-mix and commissioning currencies

The options for commissioning currencies range from a single fixed tariff payment for each episode, regardless of length of stay ('episode rates'), to payment at a daily rate ('bed-day' or '*per diem'* rates). Both systems present potential opportunities for gaming. Fixed episode payments tend to place the greater share of risk on the provider, and may encourage them to 'cream-skim' the easy cases and to discharge too early. *Per diem* payments may contain insufficient incentives to move the patient on, and so result in unnecessarily long admissions.[21]

The principal argument for episode payments is that the majority of the active treatment costs are incurred during the early part of the episode – longer-staying patients mainly incurring only 'hotel costs' due to largely avoidable delays - so restriction of payment to within a few days either side of the average length of stay provides an incentive towards efficient case throughput. However, whilst this may be true for acute medical and surgical treatments, it does not necessarily hold good for other areas of healthcare.

In surgical and acute episodes, the majority of the costs are incurred within the first few days as expected, and tail off sharply thereafter. However, in palliative care the daily costs of care rise progressively towards death, whilst in rehabilitation the cost curve is largely flat – nursing, care and therapy inputs continuing at a similar level throughout the admission. Different payment models are therefore required for these longer-term service areas.

### Case-mix in rehabilitation – international models

Rehabilitation poses some particular challenges for the development of case-mix design:

- Diagnosis alone is a relatively poor indicator of costs for in-patient rehabilitation,[36, 187] where nursing and therapy staff input (as opposed to medical treatments) are the major cost-indicators.[188-190]
- Cost-efficiency does not always equate with shorter stay. Evidence from the US and other countries has shown that the introduction of fixed episode payment schemes in rehabilitation may lead to poorer functional outcomes[191] and increased rates of discharge to institutional care[73, 192], due to pressure for early discharge when reimbursement ceases.
- Some patients need longer to achieve maximal independence, but there is also evidence that the resulting savings in the cost of on-going care can offset the initial investment in rehabilitation by several fold.[11, 32]

As a result, some healthcare systems have recognised that rigid episode-based reimbursement may be unsuitable for rehabilitation and alternative case-mix and payment models are required which are fair to both purchasers and providers and still reward efficiency.

## *The American system*

The In-patient Rehabilitation Facility Prospective Payment System uses a complex case-mix classification which was summarised concisely in a review by Richard Zorowitz.[20] Each patient is assigned into one of 85 impairment codes in 21 Rehabilitation Impairment Categories. These are further broken down into Case-mix Groups, based on the FIM Motor score on admission - and in some instances also the FIM Cognitive score and age of the patient, where these are required to provide further definition of costs. Each Case-mix Group also has four levels of co-morbidity that represent the presence or absence of medical conditions (e.g. hypertension, diabetes etc) and/or complications (e.g. tracheostomy).

Medicare also requires completion of a standardised dataset consisting of diagnostic and demographic information (the In-patient Rehabilitation Facility Patient Assessment Instrument (IRF-PAI)) to be completed for each patient within 72 hours of admission, in order to qualify for reimbursement. The

Functional Independence Measure (FIM™) is embedded in this tool. Further information may be obtained from the Centre for Medicare and Medicaid Services website: http://www.cms.gov/InpatientRehabFacPPS/

## *The Australian System*

Two separate function-related case-mix systems have been developed to classify patient episodes for different levels of reimbursement for rehabilitation.

- The CRAFT system (Case-mix Rehabilitation And Funding Tree), published in 1996 was based on 12 functional categories determined by the modified Barthel Index[38].
- The Australian National Sub-Acute and Non-Acute Patient (AN-SNAP) Classification, based on the FIM was published in 1999 by the Centre for Health Service Development (CHSD), University of Wollongong.[39]

Both systems include formal designation of services into Level 1 and Level 2, on the basis of set criteria for staffing and expertise.

The AN-SNAP classification is the more sophisticated of the two systems, and was taken up more widely across Australia and New Zealand, although states varied in the extent to which it was actually used for payment, or simply to provide information. In 2012, the Independent Hospital Pricing Authority (IHPA) was required to implement a nationally consistent activity based funding (ABF) model and the AN-SNAP classification system was selected as the ABF classification system to be used for sub acute and non-acute care, including rehabilitation.

Consecutive versions of the AN-SNAP classification for in-patient rehabilitation have become progressively more sophisticated:

1. The first version comprised 32 classes based on ten principal 'Impairment categories', subdivided by the FIM-Motor sub-scale – and, in some categories, further subdivided by the FIM-cognitive scale and age.
2. A second version (AN-SNAP-II) (2006) used the same variables, but included a total of 45 classes.
3. The current version (AN-SNAP v3) at the time of preparing this report comprises 55 classes for rehabilitation (40 overnight and 15 ambulatory). For a list of AN-SNAP v3 classes for rehabilitation, see Appendix 6.1. The classes for the impairment categories of 'stroke' and 'brain dysfunction' are provided as examples in Table 82.
4. A further update (AN-SNAP version 4) is due for release in July 2015.

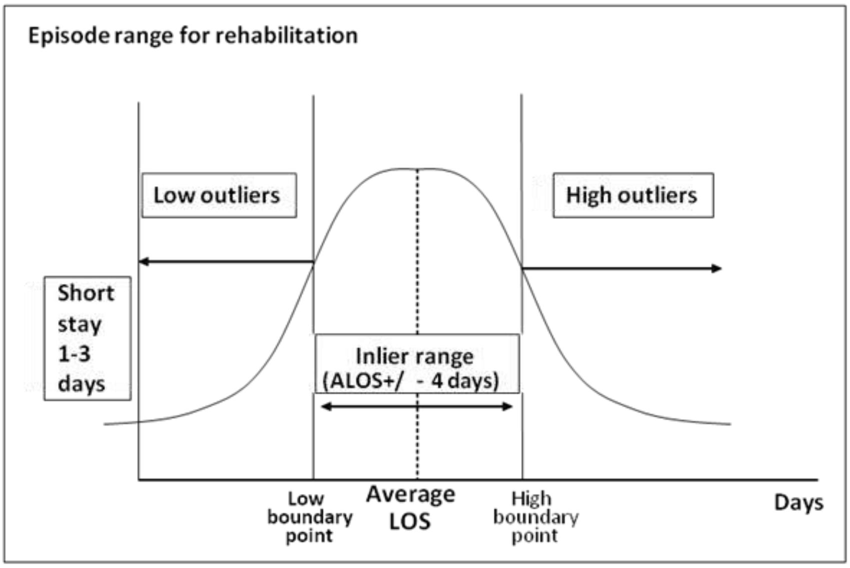
Further information may be obtained from the Australasian Rehabilitation Outcomes Centre (AROC) website: http://ahsri.uow.edu.au/aroc/index.html

**Table 82: The AN-SNAP-III case-mix classification for stroke and brain injury**

| | AN-SNAP Group | FIM score on admission | | Age |
|---|---|---|---|---|
| | | FIM Motor | FIM Cognitive | |
| | **Stroke** | | | |
| **204** | **Stroke** | 63-91 | 20-35 | |
| **205** | **Stroke** | 63-91 | 5-19 | |
| **206** | **Stroke** | 47-62 | 16-35 | |
| **207** | **Stroke** | 47-62 | 5-15 | |
| **208** | **Stroke** | 14-46 | | Age ≥ 75 yrs |
| **209** | **Stroke** | 14-46 | | Age ≤ 74 yrs |
| | **Brain Dysfunction** | | | |
| **210** | **Brain dysfunction** | 56-91 | 32-35 | |
| **211** | **Brain dysfunction** | 56-91 | 24-31 | |
| **212** | **Brain dysfunction** | 56-91 | 20-23 | |
| **213** | **Brain dysfunction** | 56-91 | 5-19 | |
| **214** | **Brain dysfunction** | 24-55 | | |
| **215** | **Brain dysfunction** | 14-23 | | |

*AN-SNAP = Australian National Sub-Acute and Non-Acute Patient classification, FIM= Functional Independence Measure*

While the AN-SNAP classification was originally developed as an episode classification, a more sophisticated 'blended payment model' for funding was subsequently developed, which is illustrated in *Figure 43*. Episode and *per diem* cost weights were derived from analysis of resource use within each of the AN-SNAP classes.[193] The model was designed to provide incentive to move patients on (as bringing in a new patient would attract a new episode weight), but still provided payment above simple hotel costs for longer episodes.




ALOS = Average Length of Stay

**Figure 43: A blended payment model**

***Legend to Figure 44:*** *Figure 44 illustrates a Blended payment model:*

- *Patients staying for very short admissions are reimbursed at a standard short-stay rate.*
- *Those below the low trim point are reimbursed at the 'Low Outlier' per diem rate.*
- *Those with lengths of stay (LOS) within the episode range attract the weighted episode rate plus an 'inlier per diem' payment for each day.*
- *Those staying beyond the episode period, attract the episode rate plus inlier per diem payment for the whole episode period, plus the days beyond the high trim point at the outlier per diem rate.*

In addition, a 90-day business rule was applied so that for patients still requiring in-patient rehabilitation beyond 90 days, the episode could be terminated and a new one started following clinical review.[193] The blended payment model has significant advantages over the more rigid US prospective payment system, as it provides flexibility for longer stay rehabilitation programmes, where clinicians and commissioners agree that these are required.

## *The InterRAI network*

The InterRAI network is an international collaborative group of researchers in over 30 countries developing case-mix systems in areas of health care for persons who are elderly, frail, or disabled. The network is not linked to any particular healthcare system or payment model, but has developed a family of instruments with the aim of providing a common language to produce integrated healthcare information across a range of settings and across international borders.

The best known of the InterRAI case-mix systems is the Resource Utilization Groups (RUG-III, RUG-IV) which is used in institutional long term care settings and skilled nursing facilities in the US and Canada. The RUG-III [194] system (version 5.12 which was current at the time of the review) uses 108 variables to create either 44 or 34 categories of patients with homogenous resource patterns. The 44-group model comprised seven main clinical groups devised as a hierarchy.

Patients were assigned to each group on the basis of clinical characteristics, and each group is divided further by an 'Activities of Daily Living' score. The Special Rehabilitation group was further divided into five levels of intensity of intervention, while patients with impaired cognition or behavioural problems were categorised on the basis of a nursing rehabilitation score.

The RUG-III algorithm explained about 55% of variance in resource use for direct patient care in nursing home residents,[195] and it has been validated in a number of countries through a series of international studies.[196-198] Further information may be obtained from the InterRAI website: http://www.interrai.org A more recent version (RUG-IV) with 66 categories is used by the US Medicare and Medicaid services in its prospective payment system for Skilled Nursing Facilities. However, these case-mix tools have not been taken up more widely in the context of specialist post-acute rehabilitation.

### Case-mix development for rehabilitation in the UK

Until 2003/4, the UK case-mix system (HRG version 3) did not include separate codes for rehabilitation. An Expert Working Group comprising case-mix developers from the (then) NHS Information Centre and clinicians from a wide range of rehabilitation service areas was set up to advise on HRG and tariff development in Rehabilitation. Case-mix development for HRGs version 4 led to the inclusion of 23 codes for rehabilitation - three for assessment and 20 for admitted inpatient care [31] (see *Box 1*).

**Box 1: Rehabilitation Healthcare Resource Groups (HRGs) version 4**

**Assessment only**

VC01Z Assessment for rehabilitation (uni-disciplinary)

VC02Z Assessment for rehabilitation (multi-disciplinary; non-specialist)

VC03Z Assessment for rehabilitation (multi-disciplinary; specialist)

**In-patient Rehabilitation (without treatment episode*)**

VC04Z Rehabilitation for stroke

VC06Z Rehabilitation for brain injuries

VC08Z Rehabilitation for spinal cord injuries

VC10Z Rehabilitation for pain syndromes

VC12Z Rehabilitation for other neurological disorders

VC14Z Rehabilitation for amputation of limb

VC16Z Rehabilitation for hip fracture

VC18Z Rehabilitation for joint replacement

VC20Z Rehabilitation for inflammatory arthritis

VC22Z Rehabilitation for non-inflammatory arthritis

VC24Z Rehabilitation for other musculoskeletal disorders

VC26Z Rehabilitation for drug and alcohol addiction

VC28Z Rehabilitation for other Psychiatric disorders

VC30Z Rehabilitation for burns

VC32Z Rehabilitation following head & neck reconstructive surgery

VC34Z Rehabilitation following other reconstructive surgery

VC36Z Rehabilitation for other trauma

VC38Z Rehabilitation for acute myocardial infarction and other cardiac disorders

VC40Z Rehabilitation for respiratory disorders

VC42Z Rehabilitation for other disorders

**(*The in-patient episode is solely for the purpose of rehabilitation, acute care episodes are coded separately)**

The classification was derived from coding data by grouper software which identified the OPCS codes for rehabilitation and ICD-10 codes for diagnosis. However, sense-checking of reports reference costs in 2005-7 demonstrated that it was not fit for purpose and that, as experience in other countries has demonstrated, diagnosis alone was not a good indicator of case-mix for rehabilitation. An initial attempt to sub-classify the HRGs by level of service was similarly unsuccessful. In the absence of any clear definitions or process for designation of Level 1, 2 and 3 services, providers classified themselves (mainly as Level 1!), and the classification still failed to differentiate on reference costs. It was subsequently accepted by the UK Department of Health that further refinement was required for the classification to be fit for the purpose of tariff development, and hence this programme was registered as a Payment by Results Improvement Project.

Preliminary exploration of the various existing international case-mix models revealed that none were immediately suitable for implementation in the UK. Case-mix classifications based on measures of physical dependency (e.g. the FIM) may work reasonably well in the context of short-stay Level 3 services, where acute stroke and orthopaedic rehabilitation form the bulk of the caseload and patients have reasonably predictable requirements for intervention. However, they work less well in specialist Level 1 and 2 programmes, where the predominant focus in the UK is on neurological rehabilitation. For this highly selected group patients with brain injury and other complex neurological disabilities, physical dependency is a much less good predictor of rehabilitation needs.

Although the FIM correlates broadly with needs for basic care in hospital and community settings,[189, 190] it does not directly measure the requirement for nursing, and correlates only weakly with the need for therapy and medical intervention.[47] It is also relatively insensitive in patients with predominantly cognitive/behavioural problems following brain injury, whose primary needs are for supervision as opposed to physical assistance.[199]

The InterRAI RUG-III system was also explored, as it includes a crude measure of intensity of intervention and an earlier investigation had shown that it had potential as a case-mix system for geriatric services in the UK.[198] However, it was found to be too limited for application in the context of complex specialised rehabilitation because of ceiling effects and, once again, the primary focus of this system is on physical dependency.

Therefore, although the well-established case-mix systems in the US and Australia have afforded valuable lessons and provided a useful model to build on, they were not in themselves fit for purpose as a case-mix and costing model for the UK.

# Collaborative work with the Australasian Rehabilitation Outcomes Centre (AROC)

AROC was established in 2002 by the Australasian Faculty of Rehabilitation Medicine, to develop a national benchmarking system to improve clinical rehabilitation outcomes in both the private and public health sectors.

- It produces information on the efficacy of interventions through the systematic collection of outcomes information in both in-patient and ambulatory settings.
- Working closely with the Centre for Health Services Development (CHSD) at Wollongong University it provides clinical and management information reports on activity and functional outcomes, classified according to the AN-SNAP 3 classification (see Chapter 2) that now form the basis of Activity Based Funding in Australia.
- It acts as the national training and accreditation centre for the Functional Independence measure (FIM™) in Australasia.

Prof Kathy Eagar, now Director of the Australian Health Services Research Institute, has led the development of AROC since its inception and is a co-applicant on this programme. Throughout this programme we have worked very closely with the AROC team, learning not only from their approaches to case-mix development and analysis, but also from their experience in running a national clinical dataset and engaging the clinical teams on the ground.

## *Comparison and case-mix and costing models*

During its development, the UKROC dataset was modelled as closely as possible to the AROC dataset to facilitate future international comparison of process and outcome data. We expected to be able to make a direct comparison of costing and outcomes between the two countries.

In the early stages of this programme we undertook an analysis of the complete AROC data for patients with neurological conditions gathered during the 10 years between 2003-2012. This has been published elsewhere and will not be presented in detail here.

This initial exploration highlighted major systems differences between the UK and Australia,

- Some 260 rehabilitation services across Australia and New Zealand currently submit data to AROC. However the large majority of these are the equivalent of Level 3 services.
- In general, the huge geographic distances over which rehabilitation services are planned in Australia, preclude the development of tertiary centres equivalent to the Level 1 services in the UK. 'Level 1' services in Australia refer to services specialising in particular diagnostic groups (e.g. brain injury) regardless of the complexity of the caseload, which would make cost-comparison non-meaningful.
- In addition the large majority of in-patient case episodes in AROC relate to orthopaedic and geriatric rehabilitation, with less than 20% for neurological conditions.

Because of these differences, it was not possible to make a direct comparison between the two countries. Instead, we performed a case-mix adjusted analysis as presented below.

## *Design*

This was a retrospective comparative cohort analysis of prospectively-collected data from the national clinical datasets for rehabilitation in the two countries, Australia and the UK.

Our aim was to compare case-mix and outcomes between England and Australia for working-aged adults admitted to in-patient rehabilitation services following acquired brain injury. We wished to determine how far the differences may be explained by case-mix and health system design, and to reflect on the adequacy of existing case-mix classifications to account for the differences between the two datasets.

## *Methods*

### Outcome measures

The principal outcome measures for the two datasets are as follows:

- AROC collects the Functional Independence Measure (FIM™)[200] version 5, recorded within 24-48 hours of admission and within 24 hours of discharge.
- UKROC dataset collects the UK Functional Assessment measure (UK FIM+FAM)[83, 132] within 10 days of admission and within 7 days of discharge.

The UK FIM+FAM incorporates the FIM version 4 (unrestricted by license), which has the same data structure as version 5. Although some small differences are recognised between the Australian and UK scoring manuals for the FIM, these do not impact significantly on the summed total or subscale scores.

### Cost-efficiency

The AROC database calculates 'FIM efficiency/week' as a proxy measure for cost-efficiency (calculated as 'FIM gain/length of stay x7'). The UKROC database uses 'time to offset the costs of rehabilitation' as its principal measure of cost-efficiency, but does collect the necessary data to calculate FIM efficiency for comparative purposes.

### Case-mix

The AROC dataset currently uses the AN-SNAP v3 case-mix classification as described above.
The UKROC database does not use the same impairment codes or AN-SNAP classes. However, it does have a diagnostic classification (comprising the site of injury and aetiology) and that can be mapped to the majority of neurological AROC Impairment codes, if completed properly. FIM data may be extracted from the UK FIM+FAM, and used to derive the same AN-SNAP classes.

### Concatenation

In both countries, patients cannot be admitted simultaneously in two hospital beds. If patients require transfer to another hospital in Australia, even if only for a single day (e.g. for imaging or investigation) this requires the patient to be discharged from the rehabilitation service and re-admitted on their return. This fragmentation of the admission into several episodes tends to yield a falsely low length of stay. Since 2012, AROC has concatenated episodes for the same patient into a single episode if the patient is re-admitted within 7 days. (Concatenation was applied retrospectively to the whole database). The UKROC dataset does not require re-admissions after a short absence to be entered as a new episode, but instead records temporary absences as 'interruptions to stay'.

### Data Extraction

Episodes were included for a 5-year period from both datasets:

- AROC: concatenated episodes for patients discharged in the calendar years 2008-2012 inclusive.
- UKROC all episodes admitted to a level 1 or 2 rehabilitation service and discharged between April 2010-March 2015.

Data were extracted for all patients aged 16-65 years, with a valid FIM rating on admission and discharge and a length of stay between 8 and 400 days (i.e. assessment admissions and long-stay episodes were excluded). Both datasets were restricted to those with an AROC Impairment code for stroke (left or right body involvement[1]) or Brain dysfunction (traumatic or non traumatic) (and the ability to derive an AN-SNAP code for Acquired brain injury (of any cause) - see Figure 44 for data extraction flow chart).

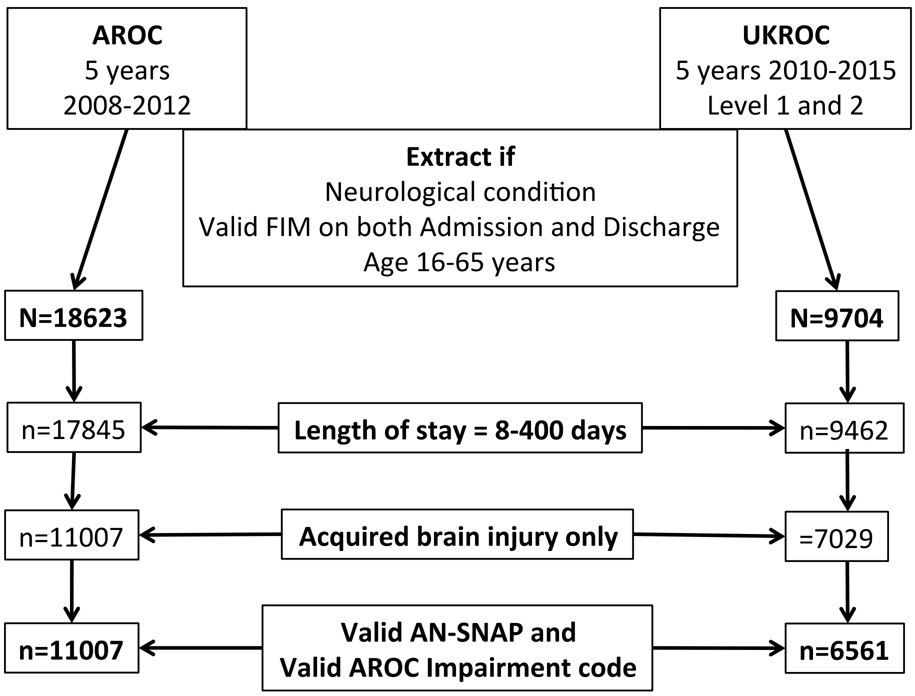



**Figure 44: Flow chart summarising data extraction from the two datasets**

[1] AROC Stroke impairment codes 1.3 (bilateral involvement), 1.4 (No paresis) and 1,.9 (other) were excluded as they could be reliably mapped from the UKROC dataset.

### Statistical analysis

Casemix adjusted outcomes were derived at case-episode level by calculating the mean figure for each AN-SNAP class and subtracting it from the actual figure within each episode.

Independent T tests were used to compare group differences for long ordinal data (i.e. ordinal scales with a large number of potential scoring levels) - such as summed subscale or total FIM scores. Mann Whitney U tests were used for comparisons of short ordinal data, such as individual item scores.
Linear regression was used to determine the extent to which the AN-SNAP classification accounted for length of stay. For this purpose, AN-SNAP classes were re-coded to an ordinal scale as follows:

- Strokes: 3-204=1, 3-205=2, 3-206=3, 3-207=4, 3-208=5, 3-209=6, 3-203=7
- Other ABI: 3-210=10, 3-211=11, 3-212=12, 3-213=13, 3-214=14, 3-215=15, 3-202=16

### Demographics

The final datasets included 11007 patients from the Australian dataset and 6561 from the UK. The demographics of the two study populations are shown in Tables 83 and 84 summarise the statistical comparisons for the main outcome parameters.

**Table 83: Demographics of the two populations**

| | Australia | | UK | |
|---|---|---|---|---|
| **Total N=** | 11007 | | 6561 | |
| **Males: Females (%)** | 64:36% | | 63:37% | |
| **Mean age: mean (SD) years** | 50.0 (13.1) | | 46.1 (13.1) | |
| **Length of stay: mean (SD) days** | 37.8 (35.9) | | 89 (70) | |
| **Diagnosis** | | | | |
| **Stroke** | 6139 (55.8%) | | 2390 (36.4%) | |
| **SAH** | 848 (7.7%) | | 645 (9.8%) | |
| **Trauma** | 2488 (22.6) | | 1296 (29.4%) | |
| **Anoxia** | 277 (2.5%) | | 510 (7.8%) | |
| **Other** | 1255 (11.4%) | | 1090 (16.6) | |
| **FIM Scores** | **Admission Mean (SD)** | **Discharge Mean (SD)** | **Admission Mean (SD)** | **Discharge Mean (SD)** |
| **FIM Motor Subscale** | 56.0 (22.7) | 78.7 (16.3) | 46.1 (27.2) | 64.5 (27.5) |
| **FIM Cognitive Subscale** | 24.1 (8.3) | 28.8 (6.1) | 19.8 (9.2) | 24.9 (8.8) |
| **FIM Total score** | 80.1 (27.0) | 107.5 (19.9) | 66.0 (33.3) | 89.4 (34.4) |
| **FIM efficiency *** | | 6.8 (5.6) | | 3.0 (3.8) |

**Individually calculated (FIM gain/ week)*

**Table 84: Statistical comparison of the two populations**

| *Parameter* | T Test* | Mean Difference | 95% Confidence Interval | | Significance |
|---|---|---|---|---|---|
| | | | Lower | Upper | (2-tailed) |
| **Age** | 19.0 | 3.9 | 3.5 | 4.3 | <0.001 |
| **Length of stay** | -55.2 | -51.0 | -52.9 | -49.2 | <0.001 |
| **FIM Motor on Admission** | 24.5 | 9.8 | 9.0 | 10.6 | <0.001 |
| **FIM Cognitive on Admission** | 30.9 | 4.3 | 4.0 | 4.6 | <0.001 |
| **FIM Total on Admission** | 29.0 | 14.1 | 13.1 | 15.1 | <0.001 |
| **FIM Motor on Discharge** | 37.9 | 14.2 | 13.4 | 14.9 | <0.001 |
| **FIM Cognitive on Discharge** | 31.6 | 3.9 | 3.7 | 4.1 | <0.001 |
| **FIM Total on Discharge** | 38.8 | 18.1 | 17.2 | 19.0 | <0.001 |
| **FIM Motor Change** | 15.4 | 4.3 | 3.8 | 4.9 | <0.001 |
| **FIM Cognitive Change** | -4.3 | -0.4 | -0.6 | -0.2 | <0.001 |
| **FIM Total Change** | 12.0 | 4.0 | 3.3 | 4.6 | <0.001 |
| **FIM Efficiency** | 54.6 | 3.9 | 3.7 | 4.0 | <0.001 |

The Australian group was older by a mean of 4 years and had a higher proportion of strokes (56% vs 36%), whilst the UK group had a higher proportion of traumatic and anoxic brain injury (37% vs 25% together).

The UK group was significantly more severely disabled than the Australian group on admission, both in motor and cognitive function, and was still more dependent by the time of discharge. UK patients also stayed in rehabilitation on average more than 50 days longer, giving an overall lower FIM efficiency. Length of stay correlated significantly (p<0.001) with total FIM scores on admission in both countries (although the relationship was slightly stronger (r=0.56) in Australia than in the UK (r=0.49).

Boxplots of the FIM subscale and total scores are shown in Figure 46. Although the overall patterns are similar, a marked tail of downward outliers is seen in the AROC group, suggesting that Australia does indeed admit a small number of severely dependent patients for rehabilitation, but that they are substantially out-numbered by the less severe group.

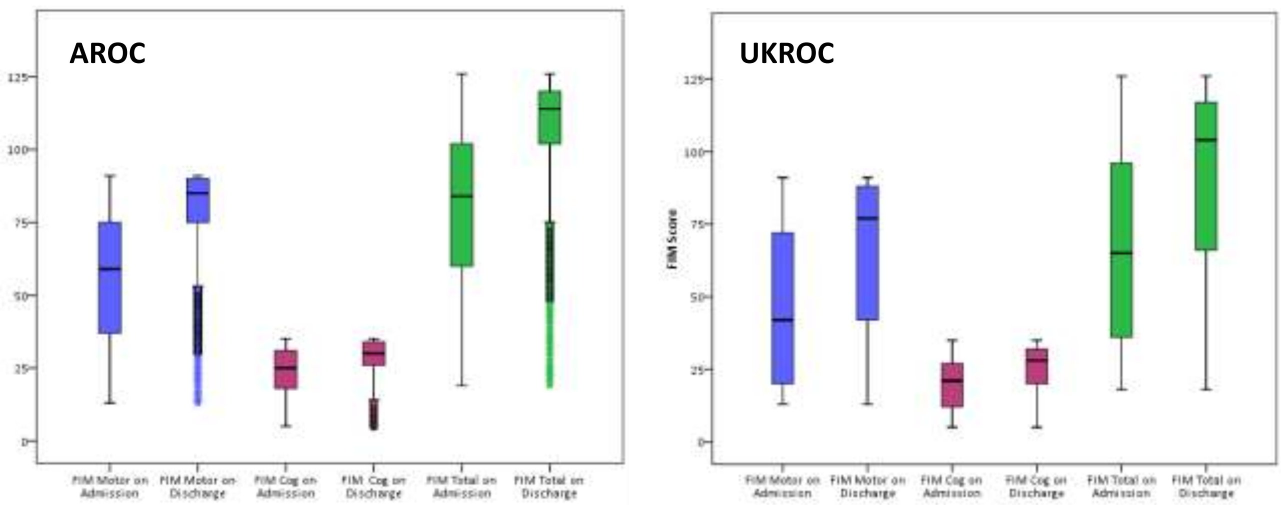



**Figure 45: Boxplot of the FIM scores on admission and discharge comparing the two countries**

We hypothesised that the UK level 2 (local specialist) services would be more similar to the Australia services Figure 46 shows the principal outcome parameters, comparing the Australian services with the UK Level 1 (tertiary) and Level 2 (local specialist) services separately. Although the differences between the two countries were less striking for the Level 2 services, they were still significantly different, so subsequent analyses were not separated by level of service.

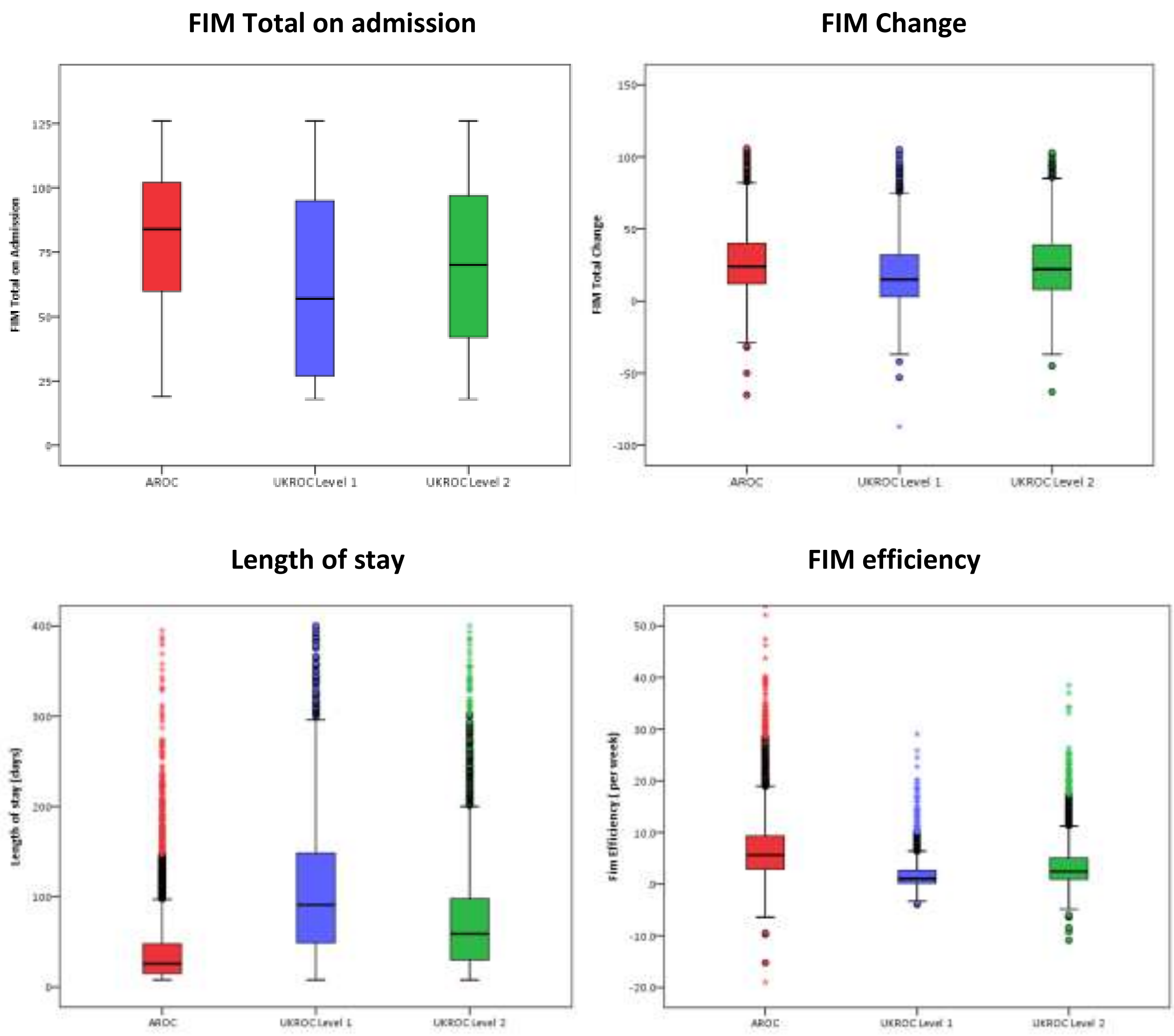



Figure 46: **Boxplots of the principal outcome parameters comparing the Australian services with the UK Level 1 and Level 2 services**

### Item level data

Figure 47 shows the FIM-Splats of the individual item scores on admission and discharge

Mann Whitney tests confirmed significantly lower scores in the UK group for all items ($z < -12.0$, $p<0.001$) except for stairs on admission ($z$ -0.56, $p=0.58$), and for all items ($z < -17.8$, $p<0.001$), including stairs ($z=-22.2$, $p<0.001$) at discharge.

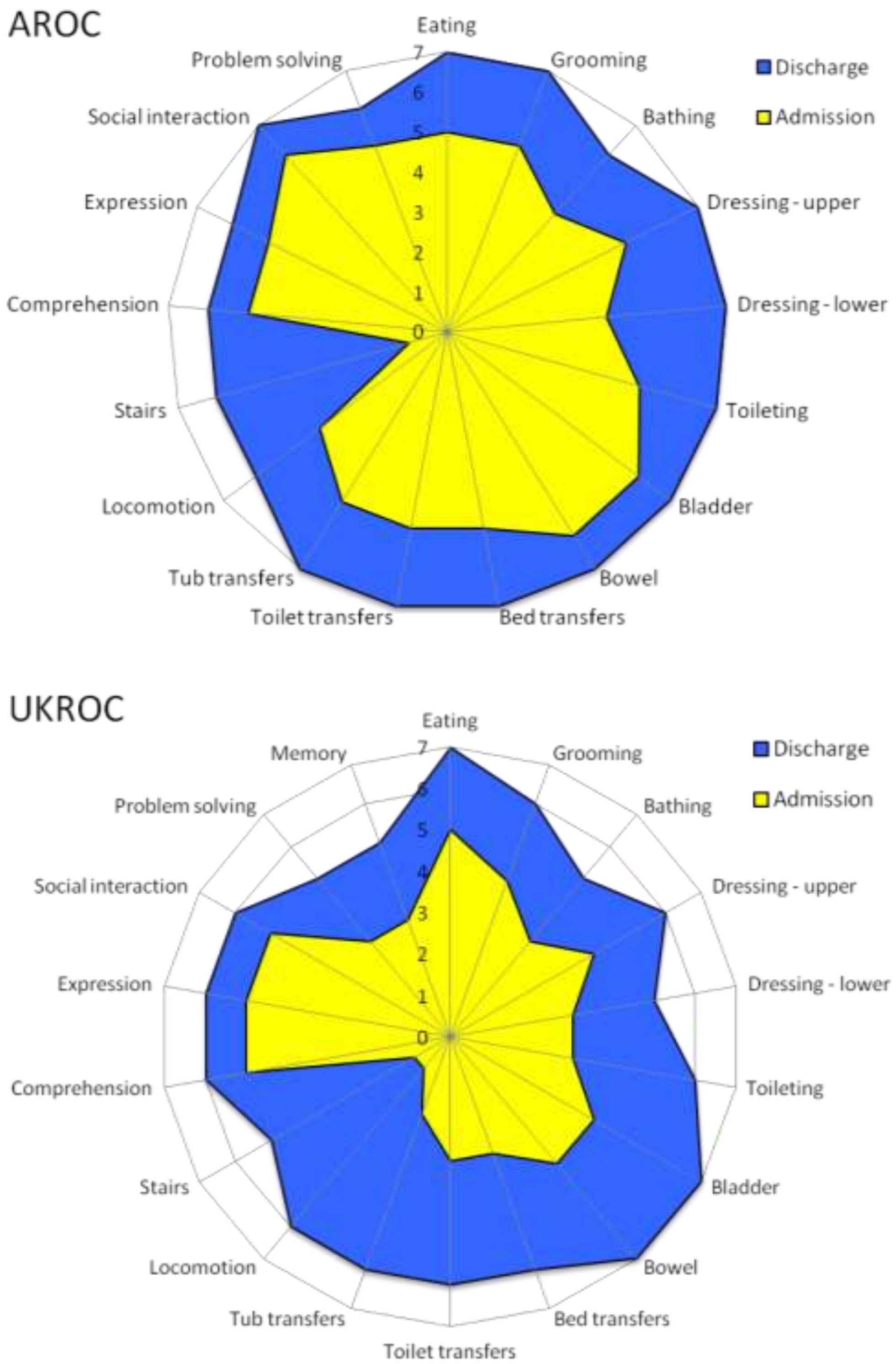



**Figure 47: FIM splats of the median admission and discharge item scores for the two populations**

### Comparison of case-mix

Case episodes from the AROC and UKROC database were classified according to their AN-SNAP class. The frequency of cases in each group is summarised in Table 85 and Figure 48. The UK sample had a greater proportion of cases in the more dependent AN-SNAP classes, especially for the non-stroke group.

**Table 85: AN-SNAP Classes version 3 and frequency of cases within each group**

| | | Australia | | UK | |
|---|---|---|---|---|---|
| **Stroke** | | N= | % | N= | % |
| 3-204 | Stroke, FIM motor 63-91, FIM cognition 20-35 | 1978 | 18% | 483 | 7% |
| 3-205 | Stroke, FIM motor 63-91, FIM cognition 5-19 | 249 | 2% | 124 | 2% |
| 3-206 | Stroke, FIM motor 47-62, FIM cognition 16-35 | 1262 | 11% | 298 | 5% |
| 3-207 | Stroke, FIM motor 47-62, FIM cognition 5-15 | 86 | 1% | 41 | 1% |
| 3-208 | Stroke, FIM motor 14-46, age>=75 | - | - | - | - |
| 3-209 | Stroke, FIM motor 14-46, age<=74 | 2461 | 22% | 1198 | 18% |
| 3-203 | All other impairments, FIM motor 13 | 103 | 1% | 246 | 4% |
| **Brain Dysfunction** | | | | | |
| 3-210 | FIM motor 56-91, FIM cognition 32-35 | 596 | 5% | 299 | 5% |
| 3-211 | FIM motor 56-91, FIM cognition 24-31 | 1276 | 12% | 776 | 12% |
| 3-212 | FIM motor 56-91, FIM cognition 20-23 | 572 | 5% | 303 | 5% |
| 3-213 | FIM motor 56-91, FIM cognition 5-19 | 764 | 7% | 406 | 6% |
| 3-214 | FIM motor 24-55 | 1177 | 11% | 1038 | 16% |
| 3-215 | FIM motor 14-23 | 358 | 3% | 611 | 9% |
| 3-202 | FIM motor 13 | 125 | 1% | 738 | 11% |

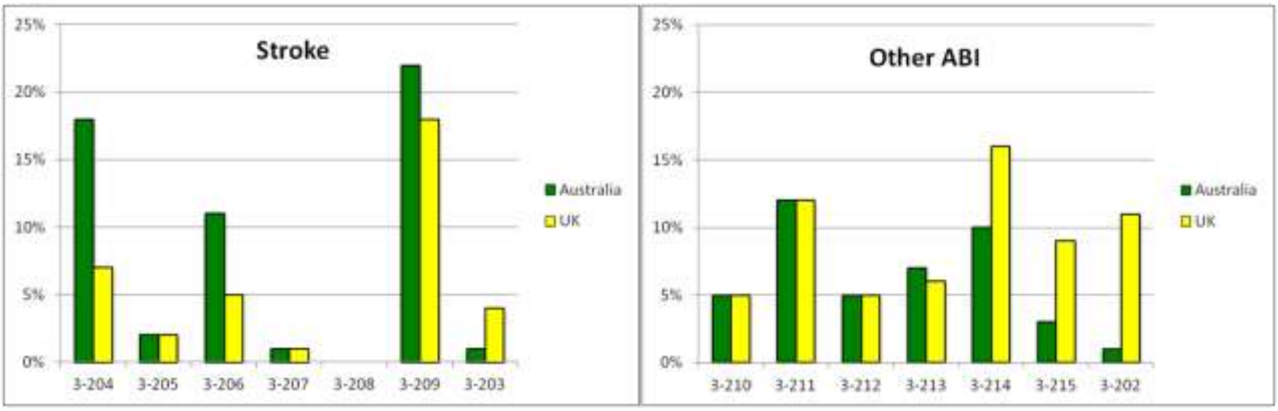



**Figure 48: Frequency distribution of AN-SNAP classes between the two countries**

Figures 49 and 50 show the FIM splats for each of the AN-SNAP classes, within each country. Within each group, the UK population tends to have a more dependent profile within the bottom two classes, (3203 and 3209 for stroke and 3214 and 3215 for brain dysfunction). The means (SD) for the principal outcomes within each AN-SNAP class for the two populations are shown in Tables 86 and 87.

Figure 49 shows the FIM splats for the various AN-SNAP classes in stroke and Figure 50 shows FIM splats for brain dysfunction.

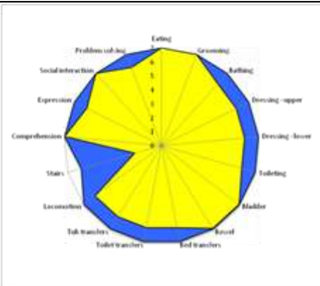
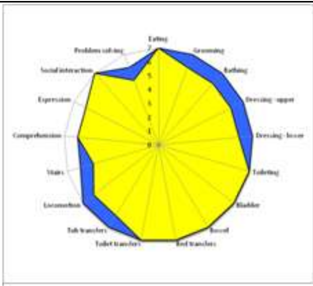
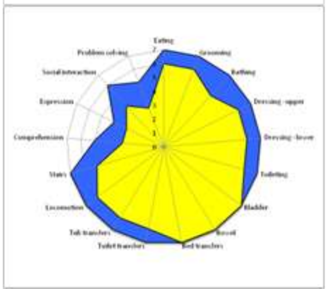
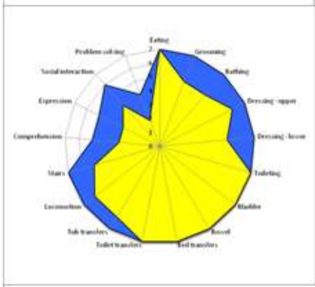
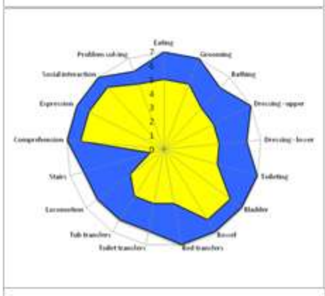
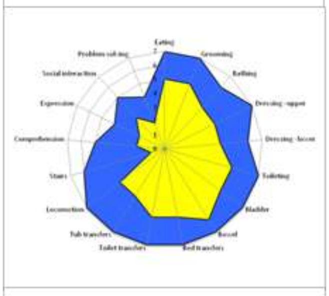
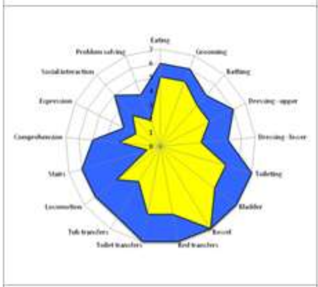
| STROKE Admission Discharge | AROC | UKROC |
|---|---|---|
| **3204** |  |  |
| **3205** |  |  |
| **3206** |  |  |
| **3207** |  |  |
| **3209** |  |  |
| **3203** |  |  |

**Figure 49:** FIM splats for the various AN-SNAP classes in stroke

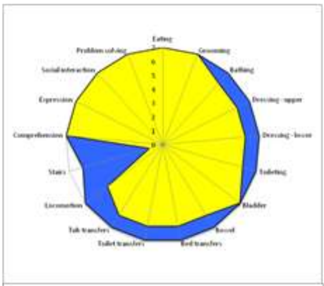
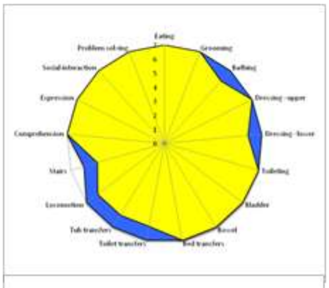
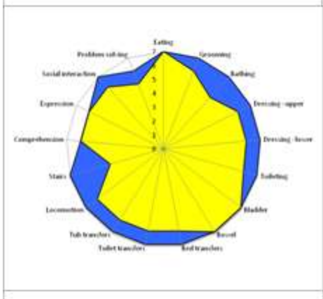
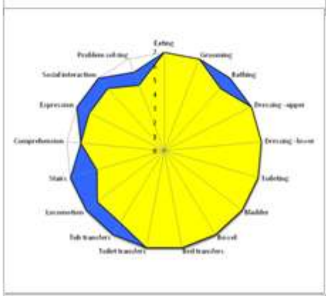
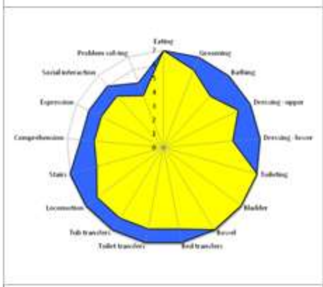
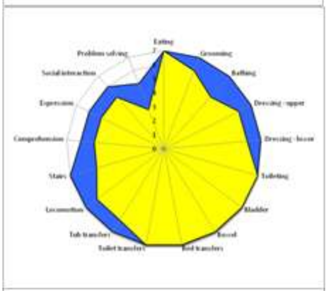
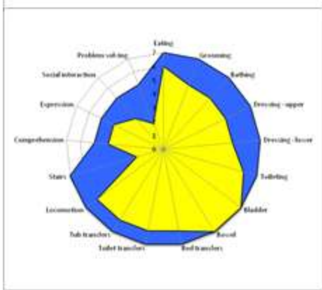
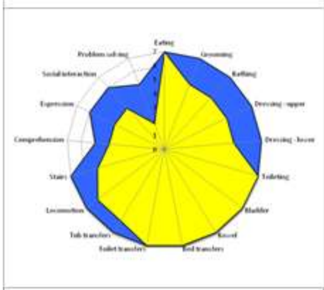
| BRAIN DYSFUNCTION Admission Discharge AROC | | UKROC |
|---|---|---|
| **3210** |  |  |
| **3211** |  |  |
| **3212** |  |  |
| **3213** |  |  |
| **3214** |  |  |
| **3215** |  |  |

**Figure 50: FIM splats for the various AN-SNAP classes in Brain Dysfunction**

**Table 86: Differences between the UK and Australian populations for the principal outcome parameters by AN-SNAP class for stroke**

| AN-SNAP class Parameter | Australia Mean | Australia SD | UK Mean | UK SD | Mean Difference | 95% CIs Upper | 95% CIs Lower | T | P value |
|---|---|---|---|---|---|---|---|---|---|
| **3-204** | **n=1978** | | **n=483** | | | | | | |
| FIM admission | 106.2 | 9.2 | 107.2 | 9.4 | -1.0 | -1.9 | -0.1 | -2.1 | 0.035 |
| FIM discharge | 118.5 | 6.4 | 116.9 | 7.1 | 1.6 | 0.9 | 2.3 | 4.5 | **<0.001** |
| FIM change | 12.4 | 8.1 | 9.8 | 8.8 | 2.6 | 1.7 | 3.5 | 5.9 | **<0.001** |
| LOS | 19.4 | 12.9 | 49.4 | 35.9 | -30.0 | -33.3 | -26.7 | -18.1 | **<0.001** |
| FIM efficiency/week | 5.5 | 4.3 | 1.9 | 2.3 | 3.5 | 3.3 | 3.8 | 25.0 | **<0.001** |
| **3-205** | **n=249** | | **n=124** | | | | | | |
| FIM admission | 91.5 | 9.0 | 91.9 | 10.5 | -0.4 | -2.6 | 1.7 | -0.4 | 0.704 |
| FIM discharge | 110.7 | 8.9 | 108.3 | 10.6 | 2.3 | 0.1 | 4.5 | 2.1 | 0.037 |
| FIM change | 19.2 | 10.0 | 16.5 | 10.7 | 2.7 | 0.5 | 5.0 | 2.4 | 0.018 |
| LOS | 29.7 | 23.4 | 86.6 | 51.1 | -56.9 | -66.4 | -47.4 | -11.8 | **<0.001** |
| FIM efficiency/week | 6.5 | 5.3 | 1.9 | 2.0 | 4.6 | 3.8 | 5.3 | 12.1 | **<0.001** |
| **3-206** | **n=1262** | | **n=298** | | | | | | |
| FIM admission | 83.2 | 7.3 | 80.9 | 6.9 | 2.3 | 1.4 | 3.2 | 5.1 | **<0.001** |
| FIM discharge | 111.7 | 11.1 | 108.0 | 10.3 | 3.6 | 2.3 | 5.0 | 5.4 | **<0.001** |
| FIM change | 28.4 | 10.7 | 27.1 | 10.3 | 1.3 | 0.0 | 2.6 | 2.0 | 0.046 |
| LOS | 28.9 | 18.3 | 66.8 | 41.6 | -37.9 | -42.8 | -33.1 | -15.4 | **<0.001** |
| FIM efficiency/week | 8.9 | 5.5 | 4.1 | 3.4 | 4.7 | 4.2 | 5.2 | 18.8 | **<0.001** |
| **3-207** | **n=86** | | **n=41** | | | | | | |
| FIM admission | 66.6 | 4.7 | 65.9 | 6.9 | 0.7 | -1.7 | 3.1 | 0.6 | 0.568 |
| FIM discharge | 102.6 | 13.9 | 96.0 | 16.1 | 6.6 | 0.7 | 12.4 | 2.2 | 0.028 |
| FIM change | 36.0 | 14.2 | 30.1 | 15.3 | 5.9 | 0.2 | 11.6 | 2.1 | 0.041 |
| LOS | 37.8 | 29.0 | 104.0 | 68.8 | -66.2 | -88.7 | -43.7 | -5.9 | **<0.001** |
| FIM efficiency/week | 9.7 | 8.1 | 3.5 | 5.1 | 6.3 | 3.9 | 8.6 | 5.3 | **<0.001** |
| **3-209** | **n=2461** | | **n=1198** | | | | | | |
| FIM admission | 52.7 | 14.7 | 48.8 | 14.0 | 4.0 | 3.0 | 4.9 | 7.9 | **<0.001** |
| FIM discharge | 93.9 | 22.2 | 81.4 | 24.5 | 12.5 | 10.8 | 14.1 | 14.9 | **<0.001** |
| FIM change | 41.1 | 19.3 | 32.6 | 19.5 | 8.5 | 7.2 | 9.9 | 12.4 | **<0.001** |
| LOS | 60.0 | 38.1 | 104.0 | 60.1 | -44.0 | -47.7 | -40.2 | -23.1 | **<0.001** |
| FIM efficiency/week | 6.4 | 4.9 | 3.1 | 3.1 | 3.3 | 3.0 | 3.6 | 24.6 | **<0.001** |
| **3-203** | **n=103** | | **n=246** | | | | | | |
| FIM admission | 26.5 | 7.2 | 23.5 | 7.1 | 3.0 | 1.3 | 4.7 | 3.6 | **<0.001** |
| FIM discharge | 66.3 | 27.3 | 39.4 | 23.9 | 26.9 | 20.8 | 33.0 | 8.7 | **<0.001** |
| FIM change | 39.9 | 27.3 | 16.0 | 21.8 | 23.9 | 18.0 | 29.9 | 7.9 | **<0.001** |
| LOS | 75.0 | 41.4 | 150.2 | 79.7 | -75.2 | -88.0 | -62.4 | -11.5 | **<0.001** |
| FIM efficiency/week | 4.2 | 3.4 | 0.9 | 1.8 | 3.3 | 2.6 | 4.0 | 9.2 | **<0.001** |

**Table 87: Differences between the UK and Australian populations for the principal outcome parameters by AN-SNAP class for other acquired brain injury**

| AN-SNAP class Parameter | Australia Mean | Australia SD | UK Mean | UK SD | Mean Difference | 95% CIs Upper | 95% CIs Lower | T | P value |
|---|---|---|---|---|---|---|---|---|---|
| **3-210** | **n=596** | | **n=299** | | | | | | |
| FIM admission | 108.4 | 9.8 | 111.6 | 10.8 | -3.2 | -4.7 | -1.8 | -4.3 | **<0.001** |
| FIM discharge | 120.9 | 5.2 | 120.3 | 6.3 | 0.6 | -0.2 | 1.5 | 1.5 | 0.129 |
| FIM change | 12.5 | 8.7 | 8.7 | 9.0 | 3.9 | 2.6 | 5.1 | 6.1 | **<0.001** |
| LOS | 17.2 | 10.7 | 31.5 | 29.0 | -14.3 | -17.7 | -10.9 | -8.3 | **<0.001** |
| FIM efficiency/week | 6.2 | 5.0 | 2.5 | 3.2 | 3.6 | 3.1 | 4.2 | 13.2 | **<0.001** |
| **3-211** | **n=1276** | | **n=776** | | | | | | |
| FIM admission | 103.7 | 11.1 | 106.8 | 11.1 | -1.5 | -3.0 | 0.0 | -1.9 | 0.053 |
| FIM discharge | 118.3 | 6.7 | 118.0 | 7.6 | 1.5 | 0.4 | 2.7 | 2.7 | 0.007 |
| FIM change | 14.6 | 10.4 | 11.2 | 10.8 | 3.0 | 1.4 | 4.6 | 3.7 | **<0.001** |
| LOS | 19.4 | 13.4 | 46.8 | 40.9 | -39.5 | -46.1 | -32.9 | -11.8 | **<0.001** |
| FIM efficiency/week | 6.4 | 5.2 | 2.7 | 3.2 | 3.5 | 3.0 | 4.1 | 12.0 | **<0.001** |
| **3-212** | **n=572** | | **n=303** | | | | | | |
| FIM admission | 96.8 | 10.9 | 98.3 | 10.7 | -1.9 | -3.2 | -0.5 | -2.7 | 0.007 |
| FIM discharge | 115.5 | 7.5 | 114.0 | 8.3 | 1.6 | 0.1 | 3.1 | 2.2 | 0.031 |
| FIM change | 18.7 | 11.4 | 15.7 | 11.6 | 3.5 | 1.7 | 5.3 | 3.8 | **<0.001** |
| LOS | 28.0 | 22.5 | 67.5 | 56.1 | -38.7 | -45.0 | -32.4 | -12.1 | **<0.001** |
| FIM efficiency/week | 6.5 | 5.4 | 2.9 | 3.3 | 3.6 | 2.9 | 4.2 | 11.4 | **<0.001** |
| **3-213** | **n=764** | | **n=406** | | | | | | |
| FIM admission | 86.9 | 11.5 | 88.7 | 11.1 | -1.9 | -3.2 | -0.5 | -2.7 | 0.007 |
| FIM discharge | 111.5 | 10.9 | 109.8 | 12.8 | 1.6 | 0.1 | 3.1 | 2.2 | 0.031 |
| FIM change | 24.6 | 14.4 | 21.1 | 15.0 | 3.5 | 1.7 | 5.3 | 3.8 | **<0.001** |
| LOS | 34.7 | 30.8 | 73.4 | 60.4 | -38.7 | -45.0 | -32.4 | -12.1 | **<0.001** |
| FIM efficiency/week | 7.3 | 6.0 | 3.7 | 4.5 | 3.6 | 2.9 | 4.2 | 11.4 | **<0.001** |
| **3-214** | **n=1177** | | **n=1038** | | | | | | |
| FIM admission | 61.4 | 13.9 | 59.7 | 13.9 | 1.7 | 0.6 | 2.9 | 3.0 | 0.003 |
| FIM discharge | 102.7 | 19.5 | 95.7 | 23.3 | 7.1 | 5.3 | 8.9 | 7.7 | **<0.001** |
| FIM change | 41.3 | 20.6 | 36.0 | 22.0 | 5.3 | 3.5 | 7.1 | 5.9 | **<0.001** |
| LOS | 46.3 | 36.8 | 84.5 | 62.5 | -38.2 | -42.5 | -33.8 | -17.2 | **<0.001** |
| FIM efficiency/week | 9.1 | 7.2 | 4.8 | 4.9 | 4.3 | 3.8 | 4.8 | 16.5 | **<0.001** |
| **3-215** | **n=358** | | **n=611** | | | | | | |
| FIM admission | 31.5 | 8.2 | 32.2 | 9.1 | -0.6 | -1.8 | 0.5 | -1.1 | 0.254 |
| FIM discharge | 82.3 | 31.7 | 67.7 | 31.7 | 14.6 | 10.5 | 18.8 | 6.9 | **<0.001** |
| FIM change | 50.8 | 31.6 | 35.5 | 29.6 | 15.3 | 11.3 | 19.3 | 7.4 | **<0.001** |
| LOS | 81.7 | 60.7 | 119.1 | 74.3 | -37.3 | -46.0 | -28.7 | -8.5 | **<0.001** |
| FIM efficiency/week | 6.3 | 6.7 | 3.2 | 4.3 | 3.1 | 2.4 | 3.9 | 8.0 | **<0.001** |
| **3-202 N=** | **n=125** | | **n=738** | | | | | | |
| FIM admission | 23.7 | 5.5 | 20.3 | 4.8 | 3.4 | 2.4 | 4.4 | 6.6 | **<0.001** |
| FIM discharge | 68.9 | 36.4 | 35.1 | 26.4 | 33.8 | 27.1 | 40.5 | 9.9 | **<0.001** |
| FIM change | 45.3 | 35.9 | 14.9 | 25.2 | 30.4 | 23.8 | 37.0 | 9.1 | **<0.001** |
| LOS | 114.1 | 81.9 | 144.2 | 85.3 | -30.2 | -45.9 | -14.5 | -3.8 | **<0.001** |
| FIM efficiency/week | 3.4 | 3.6 | 1.1 | 2.8 | 2.3 | 1.6 | 3.0 | 6.9 | **<0.001** |

Overall, within the higher functioning AN-SNAP classes, similar patterns were seen in the FIM Splats for the two countries for both the stroke and the brain injury populations. Differences in FIM scores on admission and discharge are essentially small and not clinically significant. However, in the lower functioning groups, the gains in FIM score are substantially smaller in the UK, leading to lower discharge scores for the UK sample.

Lengths of stay were significantly longer for the UK across all classes ($p<0.001$), reflecting the highly selected complex patient group treated by the UK rehabilitation services.

**Case-mix adjusted outcomes**

In order to make a direct comparison between the two countries, case-mix adjusted values for these principal parameters were computed for individual patients by subtracting the mean value for the AN-SNAP class from the patients' actual value. Thus the mean case-mix adjusted (CMA) value is expected to be zero, and the countries can be compared in relation to their mean difference around that zero point.

Table 88 shows the overall differences between the Australian and UK populations for the case-mix-adjusted main outcome parameters across all classes

**Table 88: Differences between the Australian and UK populations for the case-mix-adjusted main outcome parameters**

| OVERALL Parameter | Australia | | UK | | Mean Difference | 95% CIs | | T | P value |
|---|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | | Lower | Upper | | |
| | n=11007 | | n=6561 | | | | | | |
| FIM admission | 0.2 | 11.5 | -0.3 | 11.2 | 0.4 | 0.1 | 0.8 | 2.5 | **<0.001** |
| FIM discharge | 2.3 | 16.2 | -3.9 | 20.8 | 6.3 | 5.7 | 6.8 | 20.9 | **<0.001** |
| FIM change | 2.2 | 16.1 | -3.6 | 19.3 | 5.8 | 5.3 | 6.4 | 20.5 | **<0.001** |
| LOS | -12.3 | 30.5 | 20.6 | 61.4 | -32.8 | -34.4 | -31.3 | -40.4 | **<0.001** |
| FIM efficiency/week | 1.2 | 5.4 | -2.1 | 3.7 | 3.3 | 3.1 | 3.4 | 47.5 | **<0.001** |

*Case-mix-adjusted parameter = Actual value – mean value for the AN-SNAP class.*

The differences reach statistical significance because of the large size of the population.
Given that the AN-SNAP classification is based on admission FIM motor and cognitive ratings, it was not surprising to find that the FIM scores on admission were broadly similar. However, the case-mix adjusted gains were on average 6 points lower for the UK population, leading to lower scores on discharge. The length of stay was markedly shorter in the Australian population, leading to an overall means difference in case-mix-adjusted FIM efficiency of 3 .3 (95% CI, 3.1-3.4)

Lengths of stay, however, were substantially longer across all classes, leading to overall lower FIM efficiencies for the UK sample.

Assuming equally spaced intervals between the classes and a linear relationship between these coded values and LOS, linear regression analysis demonstrated that within the stroke subset, the ordinal AN-SNAP class accounted for 30% of the variance in length of stay for the Australian population, but only 18% of the variance in the UK population (24% overall). Within the other ABI subset, the ordinal AN-SNAP class accounted for 24% of the variance for the Australian population, and 23% in the UK population (27% overall).

## Discussion

In summary, significant differences were found, demonstrating overall that patients admitted for rehabilitation in the UK were generally more dependent on admission. Even though they stayed for longer in rehabilitation, they remained more dependent than the Australian group by the time of discharge. We expected and indeed found differences in the case-mix between the two countries. The Australian units admitted a greater proportion of strokes than the UK units. Across both diagnostic groups, a greater proportion of the UK patients were in the lower functioning groups. Thus we wanted to understand the extent to which these differences in case-mix accounted for the observed differences in length of stay and functional outcome.

Within AN-SNAP classes, the admission FIM scores were similar (as would be expected as the AN-SNAP class is principally derived from the Admission FIM). However, the change in FIM was smaller in the UK sample, giving overall lower discharge FIM scores. The differences were particularly apparent in the lower functioning case-mix classes.

Across all AN-SNAP classes the length of stay was significantly longer in the UK, by an average of more than 30 days. Linear regression modelling showed that the AN-SNAP class accounted for a greater proportion of the length of stay in Australia than in the UK. The findings suggest substantial differences in practice and/or rehabilitation populations between the two countries than cannot be explained by FIM-based case-mix alone.

## Limitations

There were some differences in the data collection processes that could potentially confound the interpretation of these data. These include:

1. The different timing of the ratings in relation to admission and discharge. However, these discrepancies are proportionate to the respective lengths of stay were not thought to contribute significantly to the group differences observed.
2. The slightly different approaches to rating of certain items (particularly bladder and bowel function) could have affected the assessments. However, inspection of the FIM splats did not reveal systematic differences for these items, and they were therefore not considered to have impacted significantly on the overall outcome scores.
3. The Australian dataset relies on concatenation to link episodes relating to each patient. If this process were incomplete it could falsely lower the length of stay. Concatenation is thought to be at least 95% successful, and the <5% linkage failures would not account for the extent of differences observed.

Thus, we do not consider that these small methodological differences are sufficient to account for the differences observed.

In addition, our use of linear regression analysis made assumptions about equal spacing and linearity of the relationship that may or may not be valid. Further exploration using regression trees is still underway at the time of preparing this report.

## Interpretation

So at first sight, the findings might suggest that that UK units may keep patients for longer because the health culture and commissioning arrangements in the UK NHS are generally more lax and support inefficient practice. But what are the differences in their respective health systems that may contribute to the observed differences between the two countries?

Firstly there are known differences in the patient populations that are recorded in the two datasets:

- The UKROC database collates data primarily from the Level 1 and 2 specialist services, which take a selected caseload of more complex patients representing <30% of the total rehabilitation population. The Level 3 services (which make up the majority of providers that contribute to the AROC database) are not represented in the UK dataset.
- Patients referred to the Level 1 and 2 UK services have, in effect, already failed Level 3 rehabilitation and one of the principal reasons for referral is the expectation that they will require longer rehabilitation programmes in a more intensive environment.
- This is particularly so for the Level 1 services. We therefore expected, and found, that the performance of the UK Level 2 services was closer to that of the Australian units, but it was still significantly different.
- We accept that a small number of Australian service providers (particularly in the major centres such as Sydney or Melbourne) probably carry a selected caseload of similar complexity to the UK units, but the dataset available for analysis did not identify individual providers, so it was not possible to make a comparison for this particular sub-group.

In theory the case-mix system should account for such differences, but may not do so adequately. The FIM scale is focused predominantly on physical function and may, therefore, be a relatively poor case-mix measure for patients with acquired brain injury who have complex rehabilitation needs relating to medical, nursing needs or and cognitive/psychosocial function are poorly captured by the FIM. This is one of the reasons why the UKROC captures a more extended dataset which includes these parameters

Alternatively, it is possible the goals for rehabilitation differ between the two countries because of the way that services are organised. When the FIM forms the basis for both case-mix and outcome measurement, and FIM-efficiency is a key parameter for national benchmarking, there is a natural tendency to focus programmes primarily on goals that are reflected in FIM scores.

The UK units do not have this restriction – in fact previous work from the UK has demonstrated that the FIM covers less than half of the domains for personal rehabilitation goals in an inpatient rehabilitation programme.[198]

- There is evidence that the provision of longer periods of rehabilitation for patients with complex needs can be highly cost-efficient despite the apparently lower FIM-efficiency in this group [11, 32]. The commitment of UK health and social services to support long term care, supports commissioning of longer stays in rehabilitation as part of evidence based clinical practice for patients who would very likely not qualify for rehabilitation programmes in Australia.
- At the other end of the scale, vocational rehabilitation / work withdrawal is quite commonly addressed in in-patient/residential services in the UK, whereas in Australia these aspects would normally be addressed in the community because of the greater distances involved.

Unfortunately it was not possible to conduct the equivalent comparative analysis on the two datasets using the UK case-mix classification and outcomes, as the large majority of Australian units do not collect the case-mix and cost-efficiency measures that underpin the UK bench-making system. A number of Australian centres did initially express interest in collecting the UK outcome measures for a sample period, but unfortunately other pressures intervened to make this impossible (see below).

In summary, we concluded that analysis of case-mix adjusted outcomes provides a useful method to facilitate cross-country comparison. Where the case-mix classification fails to account for all of the differences observed, these can then be explored further to consider whether they resonate with known differences between the respective health systems, as in this case they do.

This promotes a shared understanding of cross-cultural differences in service organisation between countries, and provides the opportunity to consider the extent to which different service models may (or may not) translate from one country to another.

## *Training and accreditation for the UK FIM+FAM*

A further area of collaboration between UKROC and AROC has been the development of training and accreditation for clinicians using the FIM+FAM. In Australia, AROC holds the national licence for the FIM™ version 5 and, under the terms of that licence, AROC is the registered centre for training and accreditation in its use. Within the centres submitting data to AROC, all staff must have received formal FIM training and must sit a credentialing exam every 2 years. This is important because the FIM is not only their outcome measure, but also their case-mix measure, which determines the rate of reimbursement for each patient.

In the UK, we have taken a somewhat different approach. The UK FIM+FAM incorporates the preceding version of the FIM (version 4) which did not require a licence from the originators. The two versions are not structurally different, but there are some subtle differences in the FIM-rating system between US, Australia and the UK (for example in the permitted use of zero scores – in which the UK and Australia are much closer than the US). In addition there are some cultural differences in the general approach to FIM assessment – the Australian manual demanding a very systematised method, breaking each task into its component steps, while the UK manual takes a more generalised approach.

Between 2008-10 we worked closely with the Australian FIM Master trainers to develop a training programme for the FAM items in Australia, and it was also our intention to extend the AROC online credentialing exam to include the FAM, in order for this to be used by both the UK and the Australian clinicians. This process led to the identification of the differences described above. In addition, in 2011 the move to Activity Based Funding in Australia sparked a major national FIM training programme, as all units were not required to submit data to AROC. Concern was raised at this point that roll out of FAM training at that time would potentially confuse clinicians.

Therefore, a decision was made not to progress online FAM credentialing at that point. Instead UKROC has continued to develop and deliver its own national training and accreditation programme for the UK FIM+FAM and other tools in the UKROC dataset to ensure their consistent application. We continue to run national workshops twice a year. We had anticipated that the demand for these training workshops would diminish as the local teams were trained in the use of the UKROC toolset. In fact, demand has continued due to staff turnover within the clinical teams. Therefore, we have focussed on developing a network of locally accredited FIM+FAM trainers through our 'Training the Trainers' programme within our UK workshops. We have also made the FIM+FAM more widely available through our website, with freely available training resources, including sets of self-service training slides for clinicians to work through in their own time and practice cases to facilitate team-based discussion.

## *In summary for this Chapter*

In this part of the programme we examined international approaches to case-mix and payment models that are already successfully running in other parts of the world, in order to assess their potential for applicability in the UK, and to learn from their advantages and limitations.

Our key deliverables were:

1. A review of international systems for case-mix and payment models in rehabilitation
2. Comparative analysis of UK and Australian datasets comparing outcomes across a range of long-term neurological conditions

3. Development of the national training and accreditation programme for the UK FIM+FAM and other tools in the UKROC dataset to ensure their consistent application.

## *Changes from the original application.*

1. We were not able to undertake a direct comparison of costs and outcomes between the Australian and UK datasets for the following reasons.

   - It became clear that the types of services contribute to the two datasets were very different – the Australian dataset consisting largely of Level 3 episodes and the UK dataset comprising a more selected sample of patients with highly complex needs.
   - Our original plans for some centres in Australia to collect the UKROC dataset were not realisable in the face of other data collection priorities imposed by the shift towards activity based funding.
   - AROC does not collect or collate information on costs.

   Nevertheless we conducted a case-mix adjusted analysis for patients with acquired brain injury using the data that the two datasets do have in common. These highlighted significant differences between the two systems, especially within the lower functioning case-mix groups.

2. We originally intended to develop online FAM credentialing using the AROC FIM credentialling system. In the event this proved impractical. Instead we continued to develop and deliver our own national training and accreditation programme for the UK FIM+FAM and other tools in the UKROC dataset to ensure their consistent application.

The programme consists of a network of local trainers together with:

- central workshops and training courses that are free to UKROC-registered teams.
- self-service FIM+FAM training tools that our freely available online.
- free technical support from the UKROC team, both for application of the tools and use of the UKROC software.

We have found this more flexible approach to be effective for the purposes of this programme, and greatly appreciated by the local clinical teams and commissioners.

# Chapter 7: Costing rehabilitation programmes in the UK

## *Part 4 Objective:*

To develop patient level-costing protocols and apply these in different specialist rehabilitation settings to determine the differential treatment costs associated with different levels of caseload complexity in the UK.

## *Key deliverables*

1. Development and implementation of a novel weighted bed-day costing and payment model for specialist rehabilitation to derive indicative tariffs for specialist rehabilitation.
2. Implementation of the model to obtain accurate costing data to underpin a set of tariffs for specialist Level 1 and 2 rehabilitation services.
3. Evaluation of the cost impact of mandating the tariffs will have been provided to NHS England

Papers describing the approaches described in this chapter have been published in Clinical Rehabilitation 2011.[131, 201]

## *Specific background*

In Chapter 6, we described the process by which the various international case-mix systems were explored, as well as those currently in use for rehabilitation in the UK. The existing HRGs were not fit for purpose as they did not distinguish the resource implications of rehabilitation for different conditions. Although the well-established case-mix systems in the US and Australia provided a useful model to build on, they were not fit for purpose as a case-mix and costing model for the UK. Instead, a different case-mix classification is proposed, based on complexity of needs for rehabilitation (including care, nursing, therapy and medical intervention) as measured by the Rehabilitation Complexity Scale[16, 47] – see Chapters 4 and 5.

The tools developed in Chapter 4 and analysis in Chapter 5 have provided detailed information about caseloads and their implications for staff time. However, under PbR, the episode tariff must reflect the full costs of treatment, which include not only other costs within the department (drugs, investigations, equipment, catering, heating, administration, maintenance etc) but also an appropriate proportion of Trust overheads (including managerial infrastructure, land and buildings etc).

In this part of the programme, we developed a standardised protocol for quantifying rehabilitation resources and obtaining patient-level costing of treatment episodes. We also describe the development of a novel multi-level weighted payment model with indicative tariffs for specialist rehabilitation. We describe a pragmatic “Patient Level Information Costing Systems” approach to costing rehabilitation services and its use to obtain accurate costing data for rehabilitation provided in different neurorehabilitation service settings for patients at different levels of complexity. Information on the cost implications of implementing the model have been provided to NHS England and Monitor to inform commissioning guidance for specialist neurorehabilitation designed to provide value-for-money services which are reimbursed at sustainable rates. The identification of ‘cost-drivers’ is expected to have wider application, not only in other areas of rehabilitation, but in other fields under the Case-mix for Long Term Care programme.  It will serve as an aid to planning and commissioning of these services.

# The UK case-mix and weighted per diem payment model

As noted in Chapter 2, the UK has a very different health culture from that in the US or Australia. The UK NHS provides the most comprehensive publicly-funded health care system in the world.  In the light of an increasingly close partnership between health and social services, rehabilitation services play a critical role, because the State is able to recoup the cost benefits of improved independence through long-term savings in the cost of ongoing care.  Under these conditions, longer lengths of stay in rehabilitation may be supported to allow patients with more complex needs to reach their full potential, providing they can be justified on grounds of cost-efficiency – i.e. it can be demonstrated that the savings in ongoing care offset the initial investment in rehabilitation.[11, 32]  However, this necessitates a different approach to payment for which the main requirements are as follows:

- Patients with complex needs are expected to incur higher treatment costs, which will vary over time.  Fair payment should be weighted in proportion to costs of providing treatment, and should also vary in relation to those costs, as the patient's needs change over time.
- The case-mix model must capture <u>all</u> the key cost-determinants, allowing for both fixed and variable costs.
- Case-mix tools must be simple and timely for clinicians to apply in routine practice.

In this section we set out the principles of the proposed UK case-mix and payment model.

- We describe the rationale of the model and the methods that are used for collection and collation of case-mix data and costing information.
- We also present preliminary data of the type that were used to underpin the development of a multi-level weighted payment model, and tariff development in the context of rehabilitation in England.

### Rationale and data collation:

The primary factors that determine the costs of treatment in rehabilitation settings are:

1. The patient's basic care and nursing and medical needs that must be provided for wherever they are managed.
2. The costs of therapy intervention and any specialist equipment and facilities required for their rehabilitation programme
3. The duration of the rehabilitation programme (i.e. length of stay in an in-patient programme).

All of these parameters are captured by the Rehabilitation Complexity Scale, but the first two are likely to change over time as the patient starts to regain independence.  Therefore, to improve the sensitivity of the case-mix classification, the weighted *per diem* payment model is based on serial complexity ratings, which are measured using the Rehabilitation Complexity Scale – currently the RCS-E v 12[48].

- The payment is weighted in proportion to the differential costs of treating patients in five bands of complexity, based on the total RCS-E score.
- The daily payment rate is adjusted according to the level of complexity, and so allows for change over time.

- Payment for the overall episode is calculated at discharge, depending on the number of days the patient spent at any given complexity level.

The critical feature of this payment model is that it is fair to both payer and provider. The provider receives reimbursement to meet the additional costs of providing for patients with complex needs. However, the payer does not continue to pay high rates for a patient who had very complex needs on admission, but who progresses to lower levels of need in the course of their recovery. Complexity may go up or down, but is expected to fall for most patients over time as they regain independence, and the corresponding reduction in payment provides an incentive towards early discharge.

In order to claim the higher rate payments for patients with complex needs, specialist services must also be able to demonstrate that they provide the additional inputs to meet those needs. Similarly, payers who meet those extra costs are entitled to ask for evidence that their investment has led to meaningful outcomes in terms of improved independence, reduced on-going care costs, or at least the attainment of individual goals for rehabilitation.

Therefore the national dataset has been established to record 'needs', 'inputs' and 'outcomes' and is collated centrally through the UKROC database, which now provides the commissioning dataset for NHS England. Service designation is contingent on registration and reporting of the full UKROC dataset for all specialist in-patient neurorehabilitation episodes.[202] RCS-E scores are gathered fortnightly for each patient and UKROC provides monthly activity reporting with automated calculation of the weighted payment for each episode from serial RCS-E ratings and length of stay data.

### A 'pragmatic PLICs' approach to costing rehabilitation services

In 2009, the Department of Health for England published standards for patient-level costing [203]. Even now, six years later, relatively few services in the UK have patient level costing systems (PLICS) sophisticated enough to allocate the direct costs of treatment to individual patients prospectively - and even if they do, they rarely include the factors necessary for accurate costing of rehabilitation services. Instead, we have taken a pragmatic approach.

Service costing information is gathered from retrospective analysis of unit budget statements and accounting costs. Reported costs are then verified by site visits to ensure that there is a consistent approach to cost definition, attribution and allocation. A standard template has been devised for attributing individual lines within the budget statement to different cost types, based on the DoH's Patient Level Costing Standards. (See Appendix 7.1). This will allow more detailed future analysis to identify the source of any large cost variances between different services.

As a starting point, costs are collated under three main cost types ('direct costs', 'indirect costs' and 'overheads') as shown in *Table 89.* The majority of direct costs vary with patient throughput or complexity, and are either 'variable' (e.g. staff pay, consumables) or 'semi-fixed' (ward and administration) costs, whilst indirect costs and overheads are largely 'fixed' costs.

**Table 89: Principal categories of service costs**

| Cost types | Definition | Examples |
|---|---|---|
| **Direct costs** (Variable / semi-fixed costs) | Costs directly associated with patient care (the running costs of the rehabilitation ward/unit. | ▪ Pay – staff salaries within the unit<br>▪ Ward consumables<br>▪ Medical supplies<br>▪ Investigations /procedures<br>▪ Running costs of equipment |
| **Indirect costs** (Fixed costs) | Contribution to central running costs (for departments and services used by patients and/or staff, which are external to the unit, but provided within the facility) | ▪ Central department costs e.g. central I/T, clinical governance, portering, medical records, etc<br>▪ Estates costs,<br>▪ Capital charges/replacement costs<br>▪ Energy / utilities |
| **Overheads** (Fixed costs) | Site-wide overheads, not directly impacting on patient care | ▪ Trust Board, Finance, Human Resources<br>▪ Public divided capital<br>▪ Repayments on any Private Finance Initiatives<br>▪ Rates |

## Bed day costs and weighted costing model

At the simplest level, average bed day costs may be derived from the total annual cost divided by the annual occupied bed days. However, these do not take account of the differential costs of treating patients with complex needs. The weighted bed-day costing model has been developed allow for this variation in cost.

In this model, patient-level weighting is applied only to the variable proportion of the bed-day cost, which is expected to vary with patient complexity. *Figure 51* shows the basis for dividing fixed and variable costs, and *Figure 52* shows the derivation of cost multipliers to convert this to a 5-tier weighted costing model.

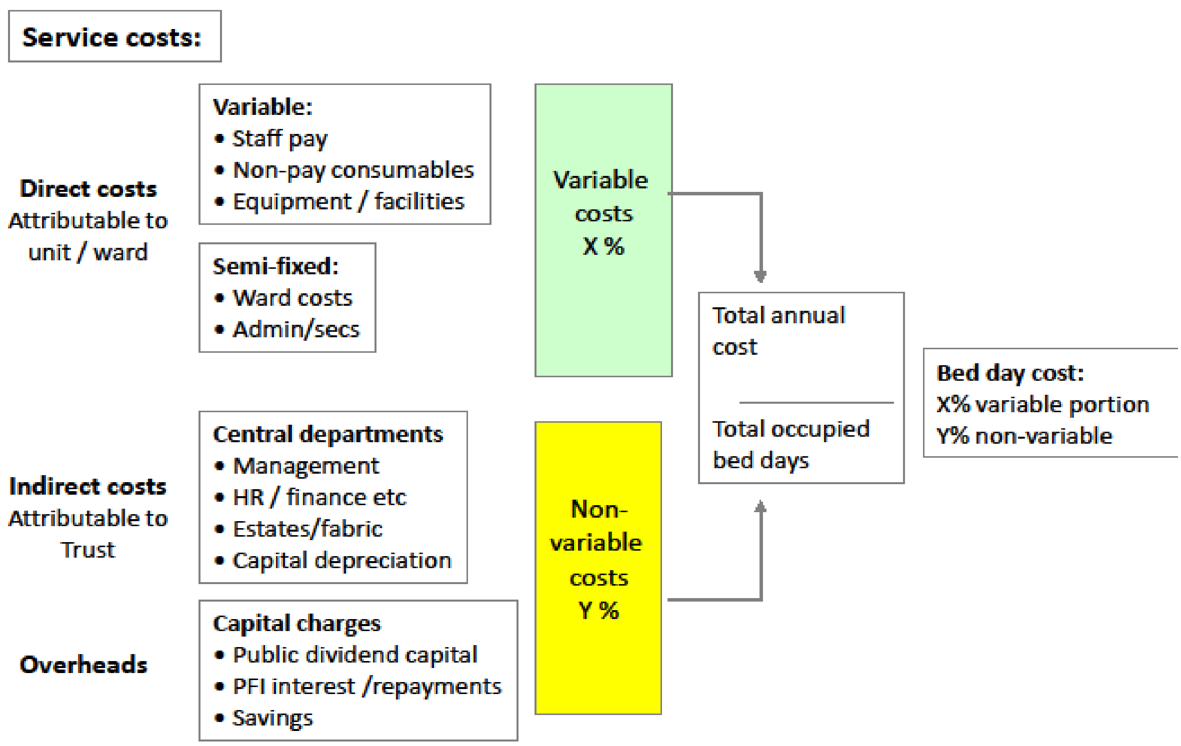



**Figure 51: Fixed and variable costs of a rehabilitation service**

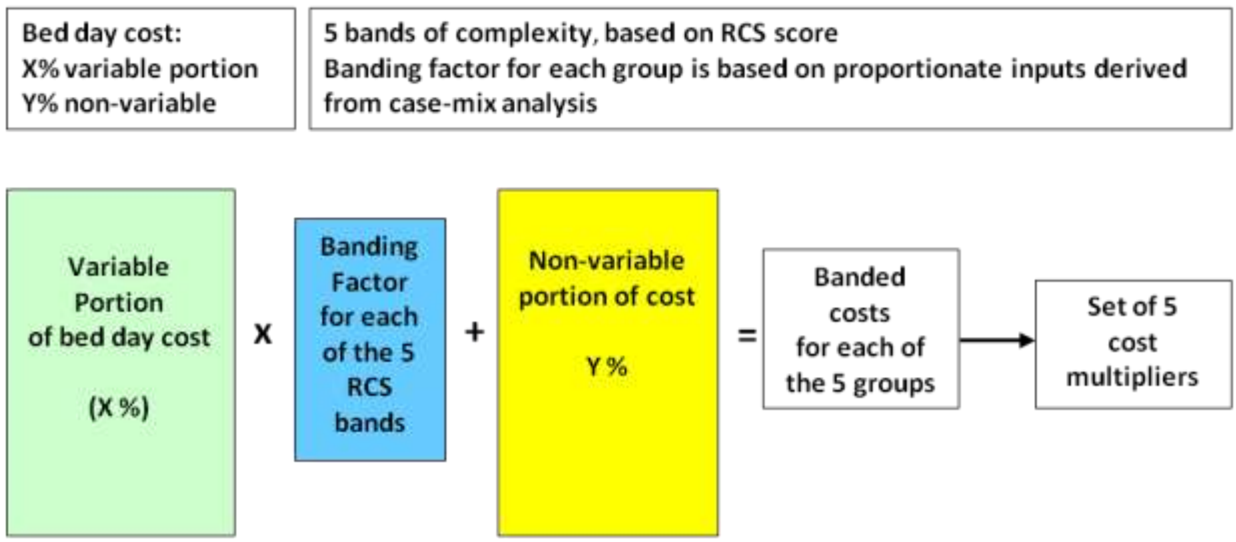



**Figure 52: Derivation of cost multipliers to produce a 5-tier weighted costing model**

As noted above, the key determinants of rehabilitation complexity are the requirements for nursing, therapy and medical care, and staff time makes up the greatest component of the variable costs. In our weighted model:

- The total Rehabilitation Complexity Scale score is used to band patients into five different levels of complexity: Very Low, Low, Medium, High, and Very High complexity, as described in Chapter 5.
- Information derived from parallel application of the Northwick Park Dependency scales is then used to determine the relative proportions of staff time associated with each complexity band (see Chapter 5).
- This staff time ratio is then applied as a banding factor to the variable portion of the bed-day cost to derive a banded cost, from which a costing multiplier is developed, as illustrated in *Figure 52*.

### Exemplar of the weighted costing model based on preliminary data collected in one service

Development of the weighted costing model is best understood through an illustrative example. The original model was developed on data from a single tertiary (Level 1) specialist inpatient neurorehabilitation service in 2010 using the RCS v8 (then the current version). Analysis of the budget statement and accounting figures from 2008/9 demonstrated that the variable costs made up approximately 75% of the total annual costs, with non-variable (fixed and semi-fixed) costs accounting for the remaining 25%.

The first step in development of the weighted costing model was to determine the relative proportions of staff time associated with each of the five complexity bands. This was undertaken on a sample of 1200 parallel ratings, see *Chapter 5, Table 66.*

The midpoint (RCS score 7-9) was taken as the reference point for the purpose of banding, and the ratio of total staff time was expressed in relation to the mean value for this Medium-complexity group, giving two higher and two lower bands.

In *Figure 53*, this staff time ratio is applied as a *'Banding factor'* to the variable portion of the bed-day cost and added to the non-variable portion to develop *banded costs*, from which a set of *costing multipliers* is derived to develop the weighted bed-day costing model. This example is set on a notional mid-point rate of £400 per bed day to simplify the calculations, but the costing multipliers so derived are independent of the figure chosen for this mid-point rate and vary only with the staff time ratio and the proportion of variable and non-variable costs.

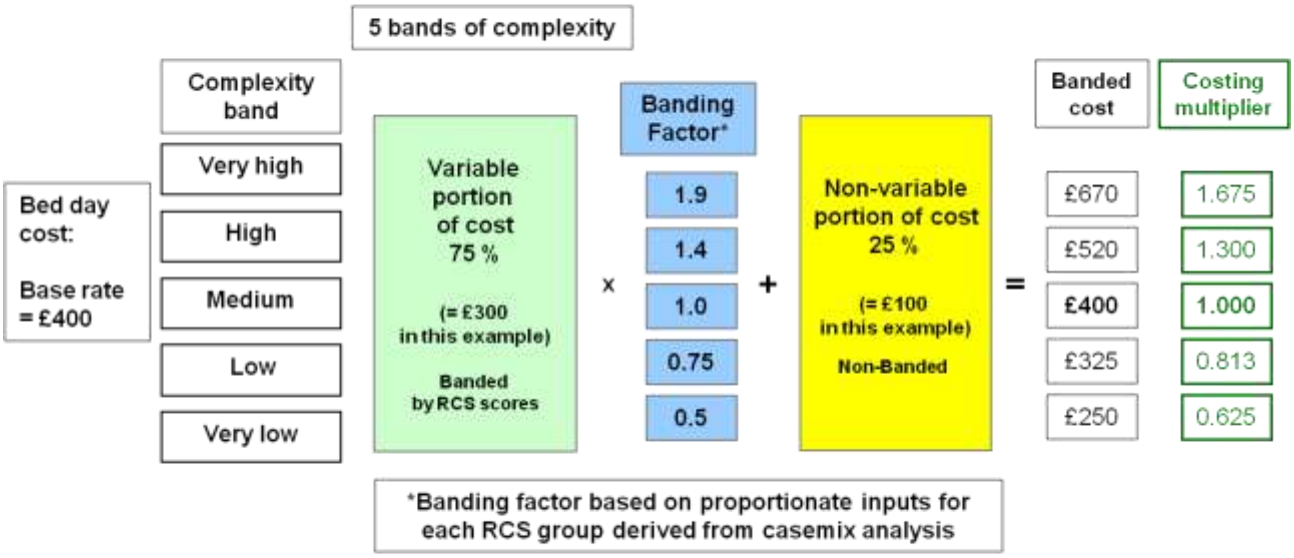



**Figure 53: Weighted costing model**

*See *Chapter 5* for derivation of the Banding factor

*Table 90* shows an exemplar calculation of how a weighted bed day cost based on this model would work out for a notional 22-bed neurorehabilitation (Level 1) service with a total annual service cost of £4m and 64% of the caseload in the 'high' or 'very high' complexity bands.

**Table 90: Example of calculation of a weighted bed day cost for a notional specialist (Level 1) in-patient neurorehabilitation service**

| **TOTAL activity** | | | **No. of weighted bed days** | | **Weighted bed day** | |
|---|---|---|---|---|---|---|
| **Complexity band** | **Costing multiplier** | **Bed days** | **X factor** | **%** | **Cost** | **Total cost** |
| Very High | 1.675 | 1,727 | 2,893 | 22% | £681 | £1,175,502 |
| High | 1.300 | 3,297 | 4,286 | 42% | £528 | £1,741,721 |
| Medium | 1.000 | 2,041 | 2,041 | 26% | £406 | £829,391 |
| Low | 0.813 | 707 | 575 | 9% | £330 | £233,575 |
| Very Low | 0.625 | 78 | 49 | 1% | £254 | £19,810 |
| **Total cost** | | **7,850** | **9,843** | **100%** | | **£4,000,000** |
| **Total annual service cost: £4,000,000.00** | | **Weighted bed day base rate** (Total annual cost/Total weighted bed days) **£406** | | | | |

** See Figure 54 for derivation of the costing multiplier.*

***Legend:*** *An example of calculation of a weighted bed day cost, taking the example of a 22-bed service with a notional bed occupancy of 7850 bed days per annum and annual cost of £4m, and a complexity profile of 64% OBDs in the High and Very High complexity bands.*

- *The Costing Multiplier (derived as in Figure 53) is applied to the occupied bed days in each complexity group to generate the number of Weighted bed days (total n=9843)*
- *The total annual cost (£4m) is then divided by the total weighted bed days, to give the Weighted Bed Day Base Rate (£406) for (i.e. the Medium complexity band).*
- *The weighted bed day cost is derived by multiplying this base rate by the costing multiplier, to resolve the total cost back to £4m*

**** Disclaimer note: this weighted cost is given as an example only. These are not actual proposed prices for commissioning.***

*Figure 54* illustrates how the multi-level payment model would be applied through serial complexity ratings


***Legend:*** *Figure 55 illustrates the application of the multi-level weighted per diem payment model in a single case. The total episode of 194 days was made up of 42 'very high', 48 'high', 70 medium and 34 'low' days, at a total cost of £92,224.*

**Figure 54: Illustration of a multi-level payment model**

This model was developed in 2009 and we recognised several limitations to the approach.

1. The data in this preliminary analysis were derived from a single centre and used the RCSv8 - which was current at the time of analysis. The model required validation in a multicentre sample using the extended version of the RCS, which addresses the ceiling effects of the RCS.

2. The estimate of staff hours provided by the Northwick Park Dependency tools may not be as accurate as formal activity analysis, but they are used here because they are timely to apply in routine clinical practice. As the introduction of Payment by Results is intended to be cost neutral, it is the relative proportion of staff time across the different bands rather than the absolute values that are of primary interest. Therefore they were considered to be acceptable for this purpose.

3. As with any case-mix system, there is the theoretical potential for gaming - for example if providers were to over-score the RCS scores to in order to maximise income. The parallel collection of dependency scores and staff profiles provides some protection against this, as inputs should be commensurate with complexity profiles, and units with low staffing levels simply could not provide the required intensity of rehabilitation (see Figure 39 in Chapter 5).

   The requirement for submission of data to the national UKROC database[204] provides a critical opportunity for benchmarking and comparison. Experience from the Australian dataset suggests that any consistent anomalies in the data are usually quite readily identified through robust central data monitoring and comparative analysis (AROC team, personal communication, 2011).

**In summary,** this section has provided the rationale for the model and provided an exemplar application. It was shared with the Rehabilitation Expert Reference Group and DoH PbR team and approved for further development to inform tariff development at a national level. The next steps were:

- To obtain actual costing data for the rehabilitation services in which it would be used.
- To examine the extent of variation between different service levels and types and so to produce a set of proposed tariffs for the different levels of service.

# Implementation of the model to obtain accurate costing data for specialist rehabilitation services

In an initial pilot study in 2010, we collated data on costing and service characteristics from 17 specialist neurorehabilitation services in England to determine the range and variation between different levels of service, and to determine the extent to which this type of cost analysis could be used to underpin tariff development for Payment by Results.[201] An abbreviated summary of this work is given below.

A total of 28 providers were approached and asked to:

1. Provide an estimation of full annual costs for their service, based on retrospective analysis of unit budget statements and accounting costs.
2. Complete a service profile questionnaire, which included information about the nature of their staffing levels, facilities, catchment population and annual activity figures.

Costs were collated using the standard costing template described above, devised to attribute individual lines within the budget statement to different cost types, in line with the UK Department of Health's Patient-level Costing Standards.[203] The template is available in electronic form - Appendix 7.1.

Each centre was provided with the template, accompanied by detailed protocols and notes to assist completion and consistency. Data returns were followed up with site visits and correspondence from expert advisors and analysts within the UKROC team, in order to clarify and verify the data and to ensure consistent approach to cost definition. Depending on the availability of information, actual costs or budget statements for 2008/9 or 2009/10 were used, and then inflated or deflated to a common price base. All data used in this section relates to the actual final figures reported by providers for the full financial year 2009-10.

### Analysis and data handling

The returns were collated in Microsoft Excel and analysed using SPSS version 18 (then current).

- Costs were divided according to principal cost type (Direct, Indirect and Overheads) and classified according to behaviour (Variable and Non-variable (fixed and semi-fixed) costs), as described above.
- Key determinants of total costs were identified using stepwise multiple regression tests.
- Descriptive statistics were calculated for principal service characteristics and costs, including range, mean and standard deviation (or median and inter-quartile range (IQR) where data were significantly skewed).
- For the purpose of comparison between the Level 1 and 2a services, costs and service characteristics were calculated per occupied bed. Mann Whitney tests were used to identify significant differences, because of the small number of services in each group.

### Results

Of the 28 units approached, 20 provided data, but three were excluded because of incomplete information. The units were distributed across England as follows: the London region (5 units), North-West (4), North-East (1), Midlands (3), South (2), South-West (2). Eight units were identified as Level 1 adult services and seven as Level 2a adult services. Two of the units were children's specialist rehabilitation services - these are shown separately for the purpose of descriptive comparison, as they differ quite markedly from the adult services (see discussion).

*Table 91* shows an analysis of costs, broken down by cost types and behaviour, across the three levels of service, as presented in the original paper. [201]

**Table 91: Analysis of services costs and breakdown of cost types between the Level 1, 2 and Children's services**

| | Level 1 – Adult Services (n=8) | | | | Level 2 – Adult Services (n=7) | | | | Children's Services (n=2) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Annual cost per occupied bed** | **Mean** | **(SD)** | **Min** | **Max** | **Mean** | **(SD)** | **Min** | **Max** | **Mean** | **(SD)** | **Min** | **Max** |
| **Direct costs** | £151,770 | (£19,794) | £121,052 | £175,268 | £122,792 | (£23,201) | £87,974 | £152,787 | £335,170 | (£50,175) | £299,691 | £370,649 |
| **Indirect costs** | £20,979 | (£9,790) | £11,635 | £41,323 | £11,852 | (£7,613) | £777 | £22,081 | £63,121 | (£5,980) | £58,893 | £67,350 |
| **Overheads** | £18,378 | (£14,766) | £5,244 | £40,299 | £11,969 | (£13,565) | £0 | £30,666 | £2,056 | (£2,907) | £0 | £4,111 |
| **Total costs** | £183,910 | (£19,863) | £150,474 | £212,821 | £146,613 | (£24,632) | £109,682 | £180,443 | £400,347 | (£41,287) | £371,152 | £429,541 |
| | **% Breakdown** | | | | **% Breakdown** | | | | **% Breakdown** | | | |
| **Variable costs** | | | | | | | | | | | | |
| **Staff pay** | 66.3 | (12.5) | 54.7 | 90.4 | 66.2 | (8.5) | 54.9 | 77.8 | 71.9 | (14.0) | 62.0 | 81.8 |
| **Non-pay** | 5.8 | (2.7) | 2.3 | 9.5 | 6.7 | (2.4) | 2.7 | 10.3 | 4.5 | (4.0) | 1.7 | 7.3 |
| **Equipment / facilities** | 0.8 | (0.6) | 0.0 | 1.5 | 0.8 | (0.7) | 0.1 | 2.0 | 2.3 | (1.7) | 1.1 | 3.5 |
| **Semi-fixed costs** | | | | | | | | | | | | |
| **Ward costs** | 9.8 | (3.2) | 5.8 | 15.5 | 8.8 | (6.0) | 0.9 | 20.3 | 4.1 | (4.9) | 0.7 | 7.6 |
| **Office admin** | 1.4 | (1.0) | 0.1 | 3.3 | 1.5 | (0.9) | 0.2 | 2.5 | 0.7 | (0.5) | 0.3 | 1.0 |
| **Total Direct** | 82.8 | (9.3) | 69.0 | 100.0 | 83.8 | (8.2) | 72.4 | 94.3 | 83.5 | (3.9) | 80.7 | 86.3 |
| **Fixed costs** | | | | | | | | | | | | |
| **Indirect** | 11.3 | (4.6) | 5.8 | 19.4 | 7.9 | (5.0) | 0.7 | 14.8 | 15.9 | (3.1) | 13.7 | 18.1 |
| **Overheads** | 9.8 | (8.1) | 3.5 | 23.0 | 8.2 | (10.0) | 0.0 | 25.7 | 0.6 | (0.8) | 0.0 | 1.1 |
| **Cost behaviour** | | | | | | | | | | | | |
| **% Variable** | 77.1 | (11.5) | 63.0 | 99.8 | 78.0 | (7.6) | 68.4 | 87.5 | 80.8 | (5.8) | 76.6 | 84.9 |
| **% Non variable** | 22.9 | (11.5) | 0.2 | 36.8 | 22.0 | (7.6) | 12.5 | 31.6 | 19.2 | (5.8) | 15.1 | 23.4 |

*SD = Standard deviation*

‘Direct’ costs made up 83.3% of the costs overall, with ‘Indirect’ costs and ‘Overheads’ making up the remaining 17%. Mann Whitney tests demonstrated no significant difference in the proportion of costs between Level 1 and 2a services. Staff pay made up 66% of the total costs of adult services for both Level 1 and 2a units. Stepwise multiple regression confirmed that Staff pay was the single strongest predictor of total costs, accounting for 95% of the variance.

Costs were also broken down into ‘ ‘Variable’ and ‘Non-variable’ components. The variable component was reasonably consistent, with a median of 75-76% across the three service types.

In order to understand the cost structure and cost-drivers of the service it was necessary to take into account the relative size and activity (occupied beds days) of the services. We anticipated that Level 1 services would in general be larger and have greater costs than the Level 2a services, due to higher levels of staffing, equipment and facilities required to support their complex caseload. *Table 92* shows an analysis of capacity, activity and staffing levels across the three service types. Unit size varied markedly (from 5-54 beds), with a mean of 23 beds (median 22, IQR 14-29) overall. As expected, Level 1 services tended to be larger than the Level 2a units, and to have a lower proportion of occupied beds, but neither trend reached statistical significance.

Staffing levels were analysed per occupied-bed (see *Table 92*). The Level 1 services had higher staffing levels overall, with significantly higher total Staff-pay costs/occupied-bed (Mann Whitney z=-2.2, p=0.03). Breakdown of staffing by discipline revealed the same trend across medical nursing and therapy disciplines, but only the therapy staffing numbers were significantly higher (z = -2.0, p = 0.05). The median total cost/occupied-bed-day was £530 (IQR 435-574) for the Level 1 services, compared with £402 (IQR 326-451) for the Level 2a services (z =-2.5, p=0.009), but was more than twice that for the children’s services (£1096, range 1017-1177).

**Table 92: Summary Descriptive statistics for capacity, activity and staffing levels compared between service levels**

| | Level 1 – Adult Services (n=8) | | | Level 2 – Adult Services (n=7) | | | Children’s Services (n=2) | |
|---|---|---|---|---|---|---|---|---|
| | Median | (IQR) | Range | Median | (IQR) | Range | Median | Range |
| No of beds | 26 | (22-34) | 15-54 | 20 | (14-28) | 12 - 30 | 9 | 5-13 |
| Average No. occupied beds | 22 | (20-26) | 14-43 | 19 | (12-25) | 12 - 27 | 7 | 4-9 |
| % occupied beds | 90% | (81-96%) | 70-99% | 96% | (90-100%) | 84-100% | 76% | 73 – 78% |
| **Staffing levels (WTE per occupied bed)** | | | | | | | | |
| Nursing staff | 1.5 | (1.4-2.4) | 1.0-2.5 | 1.8 | (1.2-2.1) | 0.9-2.2 | 3.9 | 3.7- 4.2 |
| Therapy staff | 1.1 | (0.7-1.2) | 0.6-1.3 | 0.7 | (0.6-0.7) | 0.6-0.8 | 1.1 | 0.9-1.3 |
| Medical staff | 0.14 | (0.10-.23) | 0.08-0.30 | 0.13 | (0.06-0.14) | 0.01-0.15 | 0.31 | 0.15-0.47 |
| Total staff | 2.8 | (2.4-3.6) | 1.6-3.9 | 2.3 | (2.0-2.6) | 1.6-3.0 | 5.3 | 5.2-5.4 |
| **Total cost per OBD** | **£530** | **(£435-574)** | **£412-591** | **£402** | **(£326-451)** | **£300-494** | **£1096** | **£1017-1177** |

OBD = Occupied Bed day; WTE = whole-time equivalent; SD = Standard deviation

**In summary:** This analysis confirmed the feasibility of using a pragmatic PLICS approach to estimate service costs for specialist rehabilitation services. The percentage breakdown of costs was consistent across all three service types, with Direct costs making up 83%, and the Variable component 76% overall. Staff pay made up 66% and accounted for 95% of the variance of the total costs in adult services. The cost of childrens’ services was almost twice that of the adult services. However, this latter finding should be treated with caution as it is not yet known whether these two providers are representative of children’s services in the wider NHS. A further NIHR-funded programme is in development to obtain more detail on children’s services (Programme Development Grant RP-DG 0613-10002).

In general, the greater the proportion of service costs that can be assigned directly to patient care the greater the accuracy of the final costing data.[203] The consistent finding in this study that Direct costs made up 83% suggested that the cost base was robust and could be used with some confidence to develop a costing analysis for a future tariff.

The proportion of variable and non-variable costs is required for the weighted costing model.[131] The findings that variable costs consistently made up 75-76% costs across all three service types confirmed that the 25/75% assumed in the proposed model represents a reasonable split.

International health costs analyses estimate that staff wages make up the majority of costs (65-80%) in healthcare services in most countries.[205, 206] Our findings were consistent with this. The percentage breakdown of costs was similar for Level 1 and 2a units, with staff pay accounting for two-thirds (66%) of all costs. The children's services did show a slightly higher proportion of staff costs (particularly with respect to nursing and medical staff) and equipment - reflecting the more intense general demands of caring for very disabled children, and supporting their parents - as well as the need for bespoke equipment.

As anticipated, Level 1 services cost significantly more per occupied-bed than the Level 2a services (median £530 versus £402, $p<0.01$). The difference appeared to be largely related to higher therapy staffing levels - in keeping with the requirements of a more complex caseload, and also the extended networking role that Level 1 services offer to support other local services.

**Limitations** of this study were:

- Contemporaneous cross-sectional complexity data (RCS scores) were not available from the majority of services. We were therefore not able to break down the bed-day activity by complexity group in order to apply the full weighted costing model.[131]
- We could not be certain of the validity of the reported costs or workloads. Even though reported costs were verified as far as possible by site visits and follow up correspondence from both expert advisors and analysts, the standards of financial and activity monitoring did vary somewhat between service providers.
- It was generally easier to identify costs for free-standing rehabilitation units than for those operating in a ward within a general hospital setting.

Nevertheless, some critical lessons were learned for future analyses.

- Of all the costing data collected in this analysis, staff costs were the easiest to identify. All providers stated that they had accurate information readily to hand with respect to their staffing establishment, including the details of individual salaries with on-costs for each staff member.
- In this study, staff costs accounted for 95% of the variance in total costs.
- This suggests that for future service costings, the total annual staff costs projected by 150% would provide an acceptable estimate of total service costs. The only additional information required to apply the weighted costing model would then be the *Total occupied bed days* and the *cross-sectional RCS scores*.

# Roll-out of the Pragmatic PLICS model and tariff development

Taking these lessons forward, in preparation for the transition to the new commissioning arrangements, during 2012 we contacted all registered UKROC services to request updated service profiles, and specifically to seek (a) their total staff budget costs and total occupied bed day (OBD) activity for 2011/12 and (b) their cross-sectional RCS data for 2011/12.

The costing multipliers were derived from the relative proportions of staff time within each of the five complexity bands, as presented in Chapter 5 (see *Tables 65 and 67*). The staff time ratios were slightly different between the Level 1 and 2 services, the original model providing a better fit for the Level 1 services, and the second model providing a better fit for the Level 2 services (which predominated in the sample from which the multicentre data were derived. The costing multipliers are shown in *Figure 56.*

| **Complexity** | **BED DAY COST MODELLING** | | | FROM CASEMIX ANALYSIS BASED ON 1200 RATINGS | | | |
|---|---|---|---|---|---|---|---|
| **Band** | Base cost | **25% fixed** | **75% variable** | Factor | Variable | Total cost | Costing multiplier |
| Very High | **£400** | £100 | £300 | **1.9** | £570 | **£670** | **1.675** |
| High | £400 | £100 | £300 | **1.4** | £420 | **£520** | **1.300** |
| Medium | £400 | £100 | £300 | **1** | £300 | **£400** | **1.000** |
| Low | £400 | £100 | £300 | **0.75** | £225 | **£325** | **0.813** |
| Very Low | £400 | £100 | £300 | **0.5** | £150 | **£250** | **0.625** |

**This multiplier was used for the Level 1 services**

| **Complexity** | **BED DAY COST MODELLING** | | | FROM CASEMIX ANALYSIS BASED ON 1200 RATINGS | | | |
|---|---|---|---|---|---|---|---|
| **Band** | Base cost | **25% fixed** | **75% variable** | Factor | Variable | Total cost | Costing multiplier |
| Very High | **£400** | £100 | £300 | **2** | £600 | **£700** | **1.750** |
| High | £400 | £100 | £300 | **1.5** | £450 | **£550** | **1.375** |
| Medium | £400 | £100 | £300 | **1** | £300 | **£400** | **1.000** |
| Low | £400 | £100 | £300 | **0.6** | £180 | **£280** | **0.700** |
| Very Low | £400 | £100 | £300 | **0.5** | £150 | **£250** | **0.625** |

**This multiplier was used for the Level 2 services**

**Figure 55: The weighted costing models used for Level 1 and Level 2 services**

Because the number of cases reported by each centre was too small to be treated separately, the RCS-E scores were pooled to derive the proportions for each complexity band for each service level which are shown in *Table 93.* These % were applied to the total bed day activity for each unit.

**Table 93: Proportion of cases in each RCS-E band by service level**

| **RCS Band** | **Level 1a** | | **Level 1b** | | **Level 1c** | | **Level 2a** | | **Level 2b** | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **n** | **%** | **n** | **%** | **n** | **%** | **n** | **%** | **n** | **%** |
| **Very High** | 424 | **37%** | 111 | **13%** | 3 | **3%** | 157 | 13% | 157 | 9% |
| **High** | 522 | **45%** | 505 | **58%** | 45 | **46%** | 672 | 55% | 860 | 49% |
| **Medium** | 156 | **14%** | 191 | **22%** | 40 | **40%** | 291 | 24% | 495 | 28% |
| **Low** | 37 | **3%** | 46 | **5%** | 11 | **11%** | 69 | 6% | 205 | 12% |
| **Very Low** | 16 | **1%** | 16 | **2%** | 0 | **-** | 29 | 2% | 34 | 2% |
| **Total** | **1154** | | **869** | | **99** | | **1218** | | **1751** | |

By way of illustration, *Figure 57* shows the calculations for the eight Level 2a services.

- The total cost of the service was taken as the staffing budget x 1.5
- The total occupied bed-days were derived from the activity reported to UKROC
- These two figures were entered into the Tariff Calculation Sheet shown in *Figure 56*
- The calculation sheet applies the appropriate cost multiplier for the service Level (1 or 2) to the OBD cost to derive a tariff for each complexity band.
- It also applies the cost multiplier to the OBD count to derive the WBD count. The single WBD cost is the Total cost/WBD count.

| **Tariff Calculation sheet** | | | | | | |
|---|---|---|---|---|---|---|
| | | Enter annual | | | | |
| | | Bed days | Weighted | | | |
| **RCS-E** | | **Here** | **bed days** | | % of total | |
| **Band** | **Cost Multiplier** | bed days per annum | | Tariff | bed days | Total cost |
| Very High | 1.750 | 510 | 893 | £567.24 | 13% | £289,294 |
| High | 1.375 | 2183 | 3002 | £445.69 | 55% | £972,944 |
| Standard | 1.000 | 941 | 941 | £324.14 | 24% | £305,015 |
| Low | 0.700 | 225 | 158 | £226.90 | 6% | £51,052 |
| Very Low | 0.625 | 95 | 59 | £202.59 | 2% | £19,246 |
| | | 3954 | 5052 | | | **£1,637,550** |
| | £1,637,550 | **£414.15** | **£324.14** | | | |
| | **(Enter total annual** | **OBD cost** | **WBD Cost** | | | |
| | contract cost here) | | | | | |

**Figure 56: Screenshot of individual service tariff calculation sheet using Service H as an exemplar**

WBD counts and costs were derived for each of the eight services as shown in *Figure 57.*
A set of tariffs was calculated individually for each service (*Figure 57*).

| Service | Total cost | Occupied Bed days (OBD) | OBD cost | Weighted bed days (WBD) | WBD cost |
|---|---|---|---|---|---|
| A | £1,657,286 | 4161 | £398 | 5,316 | £312 |
| B | £4,683,309 | 10882 | £430 | 13,903 | £337 |
| C | £3,284,232 | 8275 | £397 | 10,572 | £311 |
| D | £1,275,825 | 2774 | £460 | 3,544 | £360 |
| E | £4,054,067 | 10326 | £393 | 13,193 | £307 |
| F | £1,661,664 | 3123 | £532 | 3,990 | £416 |
| G | £2,039,774 | 5138 | £397 | 6,565 | £311 |
| H | £1,637,550 | 3955 | £414 | 5,053 | £324 |
| | | **Median** | **£406** | | |
| | | **Mean** | **£428** | | **£335** |

**Figure 57: Weighted bed-day costs for the eight Level 2a services**

| RCS-E Band | A | B | C | D | E | F | G | H | Mean |
|---|---|---|---|---|---|---|---|---|---|
| **Very High** | £546 | £589 | £544 | £630 | £538 | £729 | £544 | £567 | £586 |
| **High** | £429 | £463 | £427 | £495 | £423 | £573 | £427 | £446 | £460 |
| **Standard** | £312 | £337 | £311 | £360 | £307 | £416 | £311 | £324 | £335 |
| **Low** | £218 | £236 | £217 | £252 | £215 | £292 | £218 | £227 | £234 |
| **Very Low** | £195 | £211 | £194 | £225 | £192 | £260 | £194 | £203 | £209 |

**Figure 58: Individual service tariff calculation sheet for the Level 2a services**

An equivalent set of calculations was made for each of the service levels. The proposed tariffs for each service were derived from the mean of the individually applied costing models except for the Level 1a service. Here the range of costs was too great to assimilate into one cost. There was a long tail partly because some units within this group provided a hyper-acute high dependency service with around the clock medical cover to take patients directly out of the acute/intensive care services. Others provided just a 'highly physically dependent' service without acute medical cover. Therefore separate tariffs were calculated for the Level1 hyper-acute services and the remainder. Costs were not calculated for the Level1c (cognitive behavioural) services from this round as data were only available from one service in this round.

The published PbR prices are normally a 'base tariff' to which the local Market Forces Factor (MFF) payment is added individually to calculate the tariff individually for each NHS Trust. Our costing and bed day data above showed the full current cots as reported by the providers. Implicit within this is an element for MFF payment. Therefore to ensure that our approach was consistent with the general PbR approach, before publication as indicative tariffs the costs were adjusted to remove MFF using the following method.

- MFF is removed by calculating the average payment for the provider group. A base tariff total quantum of costs is thus derived.
- Tariff income by complexity band is estimated from the proposed per diem tariff. However this does not total to the known reported actual costs and so an adjuster is needed to inflate or deflate this to allow for the errors in estimation an/or differences in data used for the proportionality scores of weighting per band.
- Once this is done the MFF adjusted tariff can be again rebased to allow for the variance in total cost recovery. The MFF payment element is then also added to this.
- This provides a re-calculated tariff that allows for the severity distribution and weighting, Base costs, MFF adjustments and full cost recovery can be set.
- They were also reduced by a further 1.3% for the 'efficiency savings' which were imposed on all tariffs in that year.

As an exemplar, *Figure 59* shows the re-calculated tariffs for the Level 2b services

| Level 2b Services | | OBD | Est FYC | Cost per Day | MFF | cost ex MFF |
|---|---|---|---|---|---|---|
| C003 | I | 4161 | £1,402,257 | £337 | 1.2232 | £1,146,384 |
| C009 | J | 6546 | £2,297,646 | £351 | 1.0574 | £2,172,920 |
| C014 | K | 6588 | £2,477,088 | £376 | 1.0455 | £2,369,286 |
| C015 | L | 5238 | £2,095,200 | £400 | 1.0325 | £2,029,249 |
| C018 | M | 7052 | £2,291,900 | £325 | 1.199 | £1,911,510 |
| C025 | N | 6588 | £2,279,448 | £346 | 1.0461 | £2,178,996 |
| C036 | O | 5968 | £1,593,456 | £267 | 1.0324 | £1,543,448 |
| C057 | P | 6796 | £1,467,936 | £216 | 1.0909 | £1,345,619 |
| C071 | Q | 4471 | £1,730,277 | £387 | 1.0141 | £1,706,219 |
| C073 | R | 4380 | £1,493,580 | £341 | 1.0242 | £1,458,289 |
| C063 | S | 5891 | £2,209,125 | £375 | 1.0573 | £2,089,402 |
| | | 63679 | £21,337,913 | | | £19,951,324 |
| | AVERAGE MFF | | | | 1.069498614 | |

| Complexity Band | % | total days | per diem | tariff income | MFF adjusted per diem | Tariff adjusted for full costs | Adjusted tariff income |
|---|---|---|---|---|---|---|---|
| Very High | 9 | 5731 | £495 | £2,836,899 | £463 | £478 | £2,932,730 |
| High | 38 | 24198 | £389 | £9,413,030 | £364 | £376 | £9,731,002 |
| Medium | 33 | 21014 | £283 | £5,946,962 | £265 | £274 | £6,147,871 |
| Low | 14 | 8915 | 198 | 1765182 | 185 | £191 | 1824809.724 |
| Very Low | 6 | 3821 | 177 | 676271 | 165 | £171 | 699115 |
| Totals | | 63679 | | £20,638,364 | | | £21,335,528 |
| under recovery of cost being current less tariff income | | | | | | 699,549 | |
| under recovery as a % | | | | | | 3.2784% | |
| FINAL UNDERRECOVERY | | | | | | -2,385 | |

**Figure 59: Re-calculation adjusting for MFF in the Level 2b services**

The resultant tariffs were published in the PbR guidance for 2013/14 as shown in Figure 60.

| RCS-E Band | Level 1 HA Hyper acute | Level 1a Physical disability | Level 1b Mixed | Level 2a Supra district | Level 2b Local Specialist |
|---|---|---|---|---|---|
| Very High | £712 | £680 | £633 | £586 | £495 |
| High | £553 | £528 | £491 | £460 | £389 |
| Standard | £425 | £406 | £378 | £335 | £283 |
| Low | £346 | £330 | £307 | £234 | £198 |
| Very Low | £266 | £254 | £236 | £209 | £177 |
| Mean bed day cost | £583 | £544 | £470 | £428 | £342 |
| Weighted bed day cost | £425 | £406 | £378 | £335 | £283 |
| **Published indicative tariffs after deduction of MFF and a 1.3% cost deflator for efficiency savings** | | | | | |
| Very High | £655 | £617 | £601 | £578 | £472 |
| High | £509 | £479 | £466 | £454 | £371 |
| Standard | £391 | £368 | £358 | £331 | £270 |
| Low | £318 | £299 | £291 | £231 | £189 |
| Very Low | £245 | £231 | £224 | £206 | £169 |

**Figure 60: Calculated costs and published indicative tariffs after deductions for MFF and efficiency saving**

The above 13/14 Published Indicative tariff was based on 11/12 costs. In April 2013, the calculations were re-run using 12/13 costs along with a larger sample size and more accurate submissions from services. The result was a set of tariffs presented to Monitor in April 2013, for 14/15. No change was required except for an increase in the Level 2b tariff across all bands. The revised indicative tariffs were as shown in Table 94:

**Table 94: Revised tariffs based on the 2012/13 costs**

| RCS-E Band | Level 1HA Hyper-acute | Level 1a Physical disability | Level 1b Mixed | Level 2a Supra-district | Level 2b Local Specialist |
|---|---|---|---|---|---|
| **Very High** | £655 | £617 | £601 | £578 | £521 |
| **High** | £509 | £479 | £466 | £454 | £409 |
| **Standard** | £391 | £368 | £358 | £331 | £298 |
| **Low** | £318 | £299 | £291 | £231 | £208 |
| **Very Low** | £245 | £231 | £224 | £206 | £186 |

The introduction of PbR should be cost neutral. Table 95 shows the calculated effect on commissioning costs as a result of implementing the indicative tariffs. Although there are gains and losses within each level, the overall effect is just 0.02% of the total. These figures were provided to NHS England/Monitor in May 2013.

**Table 95:** **The effect on commissioning costs of implementing the indicative tariff 2013/14**

| Service Level | Unit no: | OBD pa from RCS reports to projected Year End 2012/13 | OBD pa from latest Service Profile | Full Year Cost for 2012/13 as reported by the Providers | Anticipated Commissioned Cost for 2013/14 applying indicative tariff to current activity profile | Anticipated effect of Tariff on Service bottom line | |
|---|---|---|---|---|---|---|---|
| **Level 1a** | 1 | 5922 | 5,953 | 2,818,774 | 2,714,430 | -104,343 | |
| | 2 | 8586 | 9,113 | 5,312,400 | 5,388,564 | 76,164 | |
| | 3 | 9940 | 11,416 | 5,285,637 | 5,460,940 | 175,303 | |
| | 4 | 7183 | 8,576 | 4,422,860 | 4,360,870 | -61,990 | |
| | 5 | 14993 | 7,096 | 4,011,960 | 3,934,346 | -77,613 | |
| | 6 | 2081 | 2,534 | 1,434,338 | 1,225,351 | -208,987 | |
| | 7 | 17318 | 9,681 | 6,022,807 | 6,472,995 | 450,188 | |
| | 8 | | 3,504 | 2,428,124 | 2,121,801 | -306,323 | |
| **Level 1a Total** | | **66023** | **57,873** | **31,736,899** | **31,679,298** | **-57,600** | **-0.18%** |
| **Level 1b** | 1 | 8023 | 8,410 | 3,842,105 | 3,931,104 | 88,999 | |
| | 2 | 5057 | 5,275 | 3,801,380 | 2,568,600 | -1,232,780 | |
| | 3 | 10953 | 10,746 | 5,113,934 | 4,433,124 | -680,810 | |
| | 4 | 6615 | 5,360 | 2,632,543 | 2,320,871 | -311,672 | |
| | 5 | 6167 | 6,295 | 3,022,704 | 3,559,260 | 536,556 | |
| | 6 | 4283 | 8,543 | 4,212,401 | 4,429,216 | 216,816 | |
| | 7 | 10620 | 8,004 | 3,159,844 | 4,377,470 | 1,217,626 | |
| | 8 | 9067 | 6,074 | 2,758,826 | 2,879,024 | 120,198 | |
| **Level 1b Total** | | **60785** | **58,707** | **28,543,734** | **£28,498,668** | **-45,067** | **-0.16%** |
| **Level 2a** | 1 | 6776 | 6,804 | 2,852,877 | 2,874,904 | 22,027 | |
| | 2 | 3636 | 3,723 | 1,740,357 | 1,542,392 | -197,965 | |
| | 3 | 7153 | 6,588 | 2,478,378 | 2,574,748 | 96,370 | |
| | 4 | 4957 | 5,110 | 2,093,744 | 2,055,029 | -38,715 | |
| | 5 | 5089 | 4,205 | 1,657,286 | 1,783,266 | 125,980 | |
| | 6 | 636 | 1,744 | 1,065,335 | 798,077 | -267,258 | |
| | 7 | 4481 | 4,114 | 1,852,149 | 1,961,676 | 109,527 | |
| | 8 | 5393 | 5,840 | 2,612,939 | 2,583,779 | -29,159 | |
| | 9 | 12533 | 13,497 | 5,086,002 | 5,699,985 | 613,983 | |
| | 10 | 6658 | 7,916 | 3,284,232 | 3,300,859 | 16,627 | |
| | 11 | 2960 | 3,696 | 1,813,491 | 1,624,886 | -188,605 | |
| | 12 | 4322 | 4,257 | 1,786,489 | 1,737,453 | -49,036 | |
| | 13 | 2947 | 3,154 | 1,213,394 | 1,232,637 | 19,243 | |
| | 14 | 5535 | 5,256 | 2,441,865 | 2,111,078 | -330,787 | |
| | 15 | | 4,748 | 1,947,320 | 2,002,571 | 55,251 | |
| | 16 | | 9,344 | 4,683,309 | 4,776,421 | 93,112 | |
| **Level 2a Total** | | **73076** | **89,996** | **38,609,166** | **38,659,762** | **50,596** | **0.13%** |
| **Level 2b** | 1 | 1853 | 2,774 | 902,051 | 970,933 | 68,882 | |
| | 2 | 4038 | 4,161 | 1,926,414 | 1,967,197 | 40,783 | |
| | 3 | 3945 | 6,935 | 1,911,398 | 2,263,453 | 352,056 | |
| | 4 | 5581 | 6,259 | 2,126,756 | 2,168,051 | 41,295 | |
| | 5 | 7200 | 7,052 | 2,319,663 | 3,210,144 | 890,481 | |
| | 6 | 6190 | 9,015 | 3,050,877 | 2,697,633 | -353,244 | |
| | 7 | 9079 | 2,190 | 1,461,518 | 871,119 | -590,399 | |
| | 8 | 7116 | 7,629 | 2,025,782 | 2,976,799 | 951,018 | |
| | 9 | 1775 | 5,470 | 1,686,009 | 2,100,012 | 414,003 | |
| | 10 | 10835 | 7,258 | 2,734,641 | 2,778,479 | 43,838 | |
| | 11 | 1491 | 3,154 | 837,095 | 1,078,579 | 241,484 | |
| | 12 | 3979 | 4,068 | 1,247,330 | 1,295,205 | 47,875 | |
| | 13 | 4823 | 4,490 | 2,273,522 | 1,549,349 | -724,173 | |
| | 14 | 6266 | 4,458 | 1,693,589 | 1,635,373 | -58,215 | |
| | 15 | 4707 | 6,368 | 2,334,990 | 2,819,467 | 484,477 | |
| | 16 | 689 | 4,205 | 1,358,354 | 1,292,971 | -65,382 | |
| | 17 | 4700 | 4,122 | 1,883,059 | 1,685,890 | -197,168 | |
| | 18 | 2903 | 2,173 | 1,038,839 | 810,257 | -228,582 | |
| | 19 | | 3,342 | 1,899,171 | 1,345,889 | -553,282 | |
| | 20 | | 6,456 | 2,484,482 | 2,446,267 | -38,215 | |
| | 21 | | 3,199 | 1,661,183 | 1,178,868 | -482,315 | |
| | 22 | | 5,144 | 1,395,869 | 1,615,284 | 219,415 | |
| | 23 | | 3,504 | 1,300,707 | 1,297,287 | -3,420 | |
| | 24 | | 3,854 | 1,830,752 | 1,355,049 | -475,703 | |
| **Level 2b Local Total** | | | **117,280** | **43,384,046** | **£43,409,557** | **25,511** | **0.06%** |
| **Grand Total** | | | **323,856** | **£142,273,845** | **142,247,285** | **-26,560** | **-0.02%** |

*OBD = Occupied Bed day, RCS = Rehabilitation Complexity Scale*

## *Updating the tariff costs in 2015*

Unfortunately, full implementation of the indicative tariffs and progression towards mandated tariffs was then held up in the transitional phase of transfer to NHSE commissioning.

The changes introduced by the Health and Social Care Act 2012, represented the most radical re-organisation in the history of the NHS to date. During the two years 2013/14 and 2014/15, a period of 'Steady-state' commissioning was permitted in order to give commissioners time to work into their new roles. Therefore, although Local Area Teams were free to take forward local initiatives, there was no absolute requirement for them to do so.

Thus, although PbR guidance for these years stated that the multi-level weighted bed commissioning currency was mandated, commissioners could chose not to use it, if they so wished. The tariffs remained indicative during this period, although local area teams were strongly encouraged to take note of them and to work towards them. Some Local Area Team commissioners chose to do so quite proactively, but others preferred to continue with their traditional contract pricing.

In the interim, further development work through the Expert Reference Panel for Rehabilitation was put on hold while Monitor worked into its new role and considered its strategy for taking forward tariff development for the future.

But meanwhile, development of the Major Trauma Networks had highlighted the critical role that the Level 1 and 2 rehabilitation services play in relieving in frontline services within the acute care pathways. NHSE and service planners around the country have called for increased capacity in specialist rehabilitation to meet increasing demands – especially for hyper-acute rehabilitation services to relieve the pressure on the Major Trauma networks.

The Clinical Reference Group (CRG) for Specialist Rehabilitation has continued to work with NHS England to refine the service specification for specialist rehabilitation services. It was recognised that the commissioning of these services came from a very low base, as specialised commissioning had been established in only three regions of the country prior to transition. The standards for service provision were thus aspirational for many parts of the country, but referred to the minimum staffing requirements for the different levels of services as set out in the BSRM Specialist Neurorehabilitation Service Standards. These have recently been updated to align with the service specification.[17]

The requirement for Level 1 and 2a services to admit a high proportion of patients with very complex (Category A) needs posed a particular challenge. Historically many units had taken a mixture of Category A and B patients to balance their caseload against their staffing numbers, and the early service cost calculations and tariffs reflected this casemix. The increase in caseload complexity required an increase in staffing establishment to ensure that even the minimum standards were met, and this meant higher service costs.

Throughout this period, therefore, UKROC recorded a progressive increase in complexity staffing and service costs. But, although it continued to report activity and costing data to the Local Area Teams and to provide advice and support for locally adjusted tariffs, there was no central mechanism by which to report the annual updates on service costs for inclusion in the PbR Guidance.

Therefore, two opposite influences operated during this period to throw the originally published tariffs out of kilter:

1. In the absence of annual costing updates that would normally have reached the central tariff pricing team from Reference Costs, the published Rehabilitation tariffs were subjected to the normal annual reductions (usually around 1.5%), but with no costing data to counterbalance these losses.

2. The actual service costs rose as providers increased their staffing levels and facilities towards the standards set out in the service specification.

In addition, not understanding fully the methodology by which the Weighted bed day costing model is applied, the central Pricing team simply applied a reduction of 1.5% to each of the 5 tariff bands. Thus, a discrepancy arose between the prices listed in the published tariffs and those generated by the UKROC costing algorithm. Small though the differences were (usually £1-2 per bed day) these amounted to not insignificant sums when multiplied over a full year's activity.

The above problems have become more evident in the commissioning round for 2015/16 for the following reasons:

- It is the first year in which steady state commissioning no longer applies.
- From April 2015, NHSE commissioning covers only the Category A activity in the Level 2a services. Previously NHSE commissioned all activity in the Level 2a services as there was no validated means for separating it. Now that this is identifiable using the PCAT Tool, Local Area Teams were required to hand the Category B activity back to the CCGs.

Accepting that NHS England has no additional resources to invest, the weighted bed day currency provides a useful mechanism for coping with this change. It supports casemix adjustment within the same bottom-line budget – essentially admitting a smaller number of more complex patients. However, as part of this adjustment it was important to re-base the cost-calculations based on current clinical practice. Since April 2015 we have started to work again with Monitor and NHSE to provide them with updated figures as set out below.

### Aims:

The purpose of this analysis was to compare the originally published tariff prices (2013/14) with a) the tariffs updated to reflect current service costs b) the tariffs required to meet the minimum requirements of the service specification.

### Methods:

Working from the data now routinely reported to the UKROC dataset, we analysed:

- staffing profiles and costs data from the latest service profiles reported in 2014/15
- activity data – weighted bed days
- complexity profiles.

The UKROC team is aware of some services that are outliers in their current designated service level. Although they have been re-signposted to commissioners they have yet to be formally re-designated. In addition a small number of services have recognised reporting anomalies that are being addressed with the provider. In order to obtain the most accurate data going forward, outliers and those with reporting anomalies were excluded from this analysis.

### Results

Table 96 shows the mean reported staffing levels for medical nursing and therapy staff for the different levels of service in comparison to the minimum recommended standards in the NHSE service specification. Figures are given in WTE per occupied bed, but multiplied up to WTE for an average 20 bed services to give an indication of the numbers if staff involved. As yet there are no published standards for the Level 1c services.

Comparing these figures with those in Table 94 (on which the tariffs published in 2013 were based), it can be seen that staffing levels have increased substantially in the last 2-3 years. In general, they are approaching (but rarely exceed) the minimum standards. Nursing staff levels fall consistently short, however – sometimes substantially so. Bearing in mind that nursing staff are responsible for the 24 hour care and safety of an increasingly unstable caseload, these shortfalls are potentially hazardous.

**Table 96: A comparison of current report mean staffing levels for the various service levels with the minimum standards recommended in the service specification**

| Service level | Mean reported | | Minimum standards | | Difference |
|---|---|---|---|---|---|
| | WTE per occupied bed | WTE per 20 beds | WTE per occupied bed | WTE per 20 beds | WTE per 20 beds |
| **Hyper-acute** | | | | | |
| Medical Staff | 0.3 | 6.8 | 0.4 | 7.0 | (0.2)* |
| Nursing Staff | 2.7 | 54.2 | 3.0 | 60.0 | (5.8) |
| Therapies Staff | 1.2 | 23.2 | 1.2 | 23.0 | 0.2 |
| **Level 1a** | | | | | |
| Medical Staff | 0.2 | 3.4 | 0.3 | 5.0 | (1.6) |
| Nursing Staff | 1.9 | 38.9 | 2.4 | 47.0 | (8.1) |
| Therapies Staff | 1.0 | 20.2 | 1.1 | 22.0 | (1.8) |
| **Level 1b** | | | | | |
| Medical Staff | 0.2 | 3.6 | 0.2 | 4.0 | (0.4) |
| Nursing Staff | 2.0 | 39.8 | 2.0 | 40.0 | (0.2) |
| Therapies Staff | 1.1 | 21.2 | 1.1 | 22.5 | (1.3) |
| **Level 1c**** | | | | | |
| Medical Staff | 0.2 | 3.5 | - | - | - |
| Nursing Staff | 2.4 | 48.9 | - | - | - |
| Therapies Staff | 1.4 | 28.5 | - | - | - |
| **Level 2a** | | | | | |
| Medical Staff | 0.2 | 3.4 | 0.2 | 4.0 | (0.6) |
| Nursing Staff | 1.7 | 33.9 | 2.0 | 40.0 | (6.1) |
| Therapies Staff | 0.9 | 18.4 | 1.0 | 20.0 | (1.6) |
| **Level 2b** | | | | | |
| Medical Staff | 0.2 | 4.0 | 0.2 | 3.0 | 1.0 |
| Nursing Staff | 1.7 | 34.8 | 2.0 | 40.0 | (5.2) |
| Therapies Staff | 0.8 | 16.4 | 0.8 | 15.3 | 1.2 |

*(* under minimum standards)* *** Based on only two services*

Table 97 shows the calculated tariffs for the different levels of service, comparing the 2013/14 tariffs with the current service costs and those that would be required to meet the requirements of the service specification.

**Table 97: Comparison of tariffs for 2013/14, with tariffs according to a) current service costs (2014/15) and b) cost of meeting minimum requirements for the service specification**

| | All prices exclude Marketing Forces Factor (MFF) | | | | | |
|---|---|---|---|---|---|---|
| **Service level** | **13/14 Prices** | **16/17 Prices Service costs** | **16/17 Prices Minimum standards** | **%** | **Price per OBD** | **Price per WBD** |
| **Hyper-acute** | | | | | | |
| Very High | £655 | £701 | £796 | 82% | | |
| High | £509 | £542 | £616 | 15% | | |
| Medium | £391 | £417 | £474 | 1% | £655 | £409 |
| Low | £318 | £338 | £384 | 0% | | |
| Very Low | £245 | £259 | £294 | 2% | | |
| **Level 1a** | | | | | | |
| Very High | £617 | £659 | £751 | 30% | | |
| High | £479 | £510 | £581 | 55% | | |
| Medium | £368 | £392 | £447 | 11% | £534 | £400 |
| Low | £299 | £318 | £362 | 2% | | |
| Very Low | £231 | £243 | £277 | 3% | | |
| **Level 1b** | | | | | | |
| Very High | £601 | £603 | £645 | 26% | | |
| High | £466 | £467 | £499 | 54% | | |
| Medium | £358 | £359 | £384 | 17% | £476 | £356 |
| Low | £291 | £291 | £311 | 1% | | |
| Very Low | £224 | £223 | £238 | 2% | | |
| **Level 1c*** | | | | | | |
| Very High | n/a | £808 | n/a | 26% | | |
| High | n/a | £625 | n/a | 55% | | |
| Medium | n/a | £481 | n/a | 15% | £636 | £479 |
| Low | n/a | £390 | n/a | 2% | | |
| Very Low | n/a | £298 | n/a | 2% | | |
| **Level 2a** | | | | | | |
| Very High | £578 | £581 | £627 | 14% | | |
| High | £454 | £457 | £492 | 51% | | |
| Medium | £331 | £332 | £358 | 26% | £432 | £345 |
| Low | £231 | £232 | £251 | 5% | | |
| Very Low | £206 | £208 | £224 | 4% | | |
| **Level 2b** | | | | | | |
| Very High | £521 | £567 | £586 | 11% | | |
| High | £409 | £446 | £461 | 52% | | |
| Medium | £298 | £324 | £335 | 26% | £407 | £340 |
| Low | £208 | £227 | £235 | 4% | | |
| Very Low | £186 | £203 | £209 | 7% | | |

*OBD = Occupied bed day, WBD = Weighted bed day*
** Based on only two services*

Table 98 summarises the cost impact for a 20-bed service based on the reported activity levels for 2014/15. The shortfall in income from the 2013/14 tariff to meet current service costs ranged from £173K in the 2a service to £329K for hyper-acute services. To the shortfall to meet the minimum requirements of the service specification, ranged from £399K to £756K.

This underlines the importance of up to date costing information in a service area that is undergoing development to meet the demands of other stages in the care pathway.

**Table 98: Comparison of activity for a 20 bed service at 2013/14 Prices with a) Current mean service costs for 2014/15 and b) Costs if meeting the minimum standards**

| Commissioning values based on the 14/15 Indicative Tariff | Service Costs Based on staffing | Service Costs Based on BSRM guidance * |
|---|---|---|
| **Level 1a Hyper-acute Service: 20 OBYs** | | |
| Commissioning value: **£4,551,480** | £4,863,729 | £5,132,101 |
| | -£312,249 | -£580,621 |
| | **-7%** | **-13%** |
| **Level 1a Service: 20 OBYs** | | |
| Commissioning value: **£3,634,577** | £3,873,106 | £4,598,618 |
| | -£238,530 | -£964,042 |
| | **-7%** | **-27%** |
| **Level 1b Service: 20 OBYs** | | |
| Commissioning value: **£3,477,854** | £3,489,269 | £3,624,163 |
| | -£11,415 | -£146,309 |
| | **-0%** | **-4%** |
| **Level 2a Service: 20 OBYs** | | |
| Commissioning value: **£3,060,862** | £3,071,688 | £3,508,700 |
| | -£10,826 | -£447,837 |
| | **-0%** | **-12%** |
| **Level 2b Service: 20 OBYs** | | |
| Commissioning value: **£2,703,317** | £2,946,279 | £3,014,668 |
| | -£242,961 | -£311,351 |
| | **-9%** | **-12%** |

* Minimum staffing provision for specialist in-patient rehabilitation service.
These recommendations are adapted from the RCP/BSRM National Guidelines for rehabilitation following Acquired Brain Injury 2003.

As yet there are no published standards for the Level 1c services. The staffing levels and costs should be regarded with some caution as the figures are based on only two services, which are quite strikingly different.

# Costing data for slow-stream rehabilitation services

It was originally anticipated that community-based rehabilitation services managing a selected group of patients with complex needs for on-going specialist nursing care or slower stream rehabilitation would be included in the NHSE service specification. However, within a short time of drawing up the initial service specifications, it became clear that the NHSE resources were already overstretched and no expansion of activity beyond the original definitions could be sanctioned.

Submission of data to UKROC was therefore continued on a voluntary basis, and some units have most helpfully provided information about staffing levels and service costs. Because data reporting was not mandated, however, the information is not as complete and robust as for the Level 1 and 2 services.

In the analysis presented below, data on staffing and costs were sought using the same methodology as for the Level 1 and 2 services. Costs were derived using the formula Total staffing budget x 1.5. Services were sub-grouped as follows:

- Services predominantly providing long-term specialist nursing care for patients with complex disabilities (n=7).
- Slow stream rehabilitation services for patients with physical disability (n=7).
- Slow stream rehabilitation services providing neurobehavioural rehabilitation for walking wounded patients with predominantly cognitive disability (n=10).

The staffing levels and costs are summarised in Table 99.

The large majority of services are provided by the Independent Sector and so do not have their own published Marketing Forces Factor (MFF). Therefore, we used the MFF for the nearest NHS Trust to calculate costs excluding MFF.

Findings of particular note are:

- As with the Level 1 and 2 services, there are economies of scale, the smaller units working out very much more expensive than the larger units.
- Specialist care homes and slow stream rehabilitation services for patients with predominantly physical disabilities report very high levels of nursing/care staff with a ratio of approximately 2.2 care assistants to 1 qualified nurse.
- The slow-stream units for walking wounded patients report high levels of therapy staff of which a variable but significant proportion are generic assistants.

**Table 99: Summary of Staffing Levels and costs for slow-stream rehabilitation/care**

| | Cost per day | Cost per day | | Staffing WTE per bed | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Gross £ | less MFF £ | Bed base | Medical | Care Assistants | Qualified Nursing | Total Nursing+Care | Therapy assistants | Qualified Therapists | Total therapy staff |
| **Predominantly Care** | | | | | | | | | | |
| **1** | £251 | £221 | 50 | 0.02 | 1.28 | 0.59 | 1.86 | 0.22 | 0.01 | 0.23 |
| **2** | £255 | £241 | 34 | 0.00 | 2.05 | 0.45 | 2.50 | 0.18 | 0.01 | 0.19 |
| **3** | £308 | £300 | 26 | 0.02 | 1.16 | 0.74 | 1.90 | 0.00 | 0.35 | 0.35 |
| **4** | £305 | £295 | 8 | 0.03 | 1.35 | 0.62 | 1.97 | 0.13 | 0.24 | 0.37 |
| **5** | £195 | £162 | 6 | 0.02 | 0.87 | 0.82 | 1.68 | 0.00 | 0.17 | 0.17 |
| **6** | £339 | £316 | 14 | 0.04 | 1.62 | 0.66 | 2.28 | 0.07 | 0.18 | 0.26 |
| **Mean (SD)** | **£76 (52)** | **£256 (59)** | **23 (17)** | **0.02 (0.01)** | **1.39 (0.41)** | **0.64 (0.12)** | **2.03 (0.30)** | 0.10 (0.09) | 0.16 (0.13) | 0.26 (0.08) |
| **Median** | **£280** | **£268** | **20** | **0.02** | **1.31** | **0.64** | **1.93** | **0.10** | **0.18** | **0.25** |
| **Slow-stream Rehabilitation - Physical disability** | | | | | | | | | | |
| **1** | £363 | £363 | 12 | 0.08 | 0.98 | 1.19 | 2.16 | 0.00 | 0.61 | 0.61 |
| **2** | £413 | £385 | 10 | 0.03 | 1.82 | 0.61 | 2.42 | 0.04 | 0.80 | 0.84 |
| **3** | £309 | £272 | 60 | 0.01 | 2.56 | 0.41 | 2.97 | 0.17 | 0.01 | 0.18 |
| **4** | £281 | £241 | 10 | 0.06 | 0.91 | 0.10 | 1.01 | 0.10 | 0.69 | 0.79 |
| **5** | £570 | £552 | 9 | 0.11 | 1.18 | 1.22 | 2.40 | 0.00 | 1.20 | 1.20 |
| **6** | £405 | £337 | 9 | 0.07 | 0.87 | 0.58 | 1.46 | 0.00 | 0.37 | 0.37 |
| **7** | £335 | £324 | 15 | 0.03 | 2.03 | 0.61 | 2.64 | 0.07 | 0.32 | 0.40 |
| **8** | £372 | £259 | 9 | 0.05 | 1.51 | 0.75 | 2.26 | 0.00 | 0.19 | 0.19 |
| **Mean (SD)** | **£381 (88)** | **£342 (99)** | **17 (18)** | **0.05 (0.03)** | **1.48 (0.61)** | **0.68 (0.37)** | **2.17 (0.64)** | 0.051 (0.06) | 0.52 (0.38) | 0.57 (0.36) |
| **Median** | **£368** | **£330** | **10** | **0.05** | **1.34** | **0.61** | **2.33** | **0.02** | **0.49** | **0.51** |
| **Slow-stream Rehabilitation - Walking Wounded** | | | | | | | | | | |
| **1** | £269 | £265 | 16 | 0.01 | 0.00 | 0.07 | 0.07 | 1.74 | 0.59 | 2.33 |
| **2** | £357 | £322 | 8 | 0.01 | 0.00 | 0.08 | 0.08 | 1.94 | 0.86 | 2.80 |
| **3** | £334 | £287 | 18 | 0.01 | 0.00 | 0.06 | 0.06 | 1.16 | 0.48 | 1.65 |
| **4** | £283 | £269 | 19 | 0.01 | 0.00 | 0.06 | 0.06 | 1.70 | 0.54 | 2.24 |
| **5** | £255 | £235 | 24 | 0.01 | 0.00 | 0.05 | 0.05 | 1.53 | 0.40 | 1.93 |
| **6** | £169 | £162 | 23 | 0.01 | 0.00 | 0.05 | 0.05 | 1.49 | 0.23 | 1.72 |
| **7** | £327 | £311 | 15 | 0.01 | 0.00 | 0.08 | 0.08 | 1.55 | 0.65 | 2.20 |
| **8** | £386 | £373 | 34 | 0.03 | 0.64 | 0.00 | 0.64 | 1.42 | 0.69 | 2.10 |
| **9** | £317 | £289 | 17 | 0.06 | 0.00 | 0.45 | 0.45 | 0.00 | 0.55 | 0.55 |
| **10** | £402 | £396 | 11 | 0.11 | 0.00 | 0.22 | 0.22 | 1.21 | 0.50 | 1.71 |
| **Mean (SD)** | **£310 (69)** | **(291 (67)** | **19 (7)** | **0.03 (0.03)** | **0.06 (0.2)** | **0.11 (0.13)** | **0.17 (0.21)** | 1.37 (0.54) | 0.55 (0.17) | **1.92 (0.60)** |
| **Median** | **£322** | **£288** | **17** | **0.01** | **0.00** | **0.07** | **0.07** | **1.51** | **0.55** | **2.02** |

Following acquired brain injury, cognitive /behavioural problems tend to improve along a somewhat slower trajectory than physical problems.  Neurobehavioural rehabilitation therefore takes a number of different forms including:

a) Highly specialist subacute neurobehavioural rehabilitation as patients go through a agitated and aggressive stage – and for example may need treatment under section of the Mental Health Act.

b) Specialist neurobehavioural rehabilitation in neuro-psychiatry-led services to manage patients with highly aggressive / challenging behaviours.

c) Slow stream cognitive rehabilitation, - services for patients who are generally on their feet ('walking wounded') but may be confused and disorientated - albeit more compliant than patients in group a) or b) services.

Within this programme we have obtained limited data costing data on services in categories a) ad c) above, but not really on services in category b) which tend to fall under the mental health tariffs and do not submit data to UKROC.

## *Costing data: Out-reach services*

Patients with complex disabling conditions do not travel easily to out-patient appointments. It is often more appropriate for consultant specialist in rehabilitation medicine (and/or their members of the multidisciplinary rehabilitation team) to visit patients in their own setting to provide assessment and/or treatment, rather than asking the patient to travel in to the centre.

Outreach services also provide a useful opportunity for the consultant in RM to communicate with the family and local treating teams, providing an opportunity both to glean useful information about the patient and to offer education and advice towards managing patients in the community and so improving the overall quality of care. This can be particularly useful for patients with very complex needs or challenging family situations were a case conference held on site as part of the assessment can sometimes help to share information transparently and defuse a difficult situation.

Outreach services are also increasingly in demand as part of the acute trauma and neurosciences pathways, to draw up the Rehabilitation Prescription, and to signpost the patient down the appropriate part of the care pathway. The CRG Service Specification for Major Trauma requires all patients with Injury Severity Scores of >8 to be seen by a consultant in rehabilitation Medicine within 96 hours. In addition, the British Society of Rehabilitation Medicine has set out core standards for delivery of the Specialist Rehabilitation Prescription within the within the Major Trauma and Acute Care Pathways.[122]

As it is not expected that every Major Trauma Centre will have a consultant rehabilitation physician on site, this service may be delivered appropriately through Outreach services from the Level 1 and 2 specialist rehabilitation services. It was therefore pertinent to develop costing to underpin tariffs for commissioning of these outreach services.

Level 1 and 2 specialist rehabilitation services are relatively thinly spread around the country. Providing an out-reach service to a large catchment area can be quite a costly exercise, although there is considerable potential for cost savings if a proportion of outreach visits could avoid an inappropriate in-patient admission.

At the outset of this programme there were no published tariffs for outreach, so the following exercise was undertaken to develop a set of tariffs for this type of work and to estimate potential cost savings. We are indebted to Dr Charlie Nyein for coordinating this study.

## Method

Initial focus group discussions with specialist rehabilitation teams that regularly perform outreach visits identified the following components to cost:

1. Clinical activity time for the assessment itself (undertaken by a Consultant in RM, or other appropriately experienced members of the specialist rehabilitation team, e.g. an Associate Specialist doctor in RM or consultant allied health professional (AHP)
2. Time for liaison with the family / team/ review of records
3. Time for preparing a report, follow-up phone calls, ongoing negotiation with commissioners, family etc
4. Travelling time – depending on distance
5. Secretarial time for making appointment, notifying parties, typing and distributing reports etc

Travel time was considered separately from clinical time, accepting that services cover widely varying catchment areas and that sometimes multiple assessments may be carried out during one trip (e.g. to a specialist nursing home).

A costing matrix was therefore drawn up to reflect four categories of assessment and four categories of travel time.

***Clinical Activity Time***

- **Standard** – Review of medical records, assessment/treatment, and discussion with patient and/or family and or treating team (**Up to 1 hour in total**)
- **Extended** – Review of medical records, assessment/treatment, and discussion with patient and/or family and or treating team (**Up to 2 hours in total**)
- **Complex** – Review of medical records, assessment/treatment, discussion with patient and/or family and or treating team + at least one other activity (**Up to 4 hours in total**)
- **Highly Complex** – Review of medical records, assessment/treatment, discussion with patient and/or family and or treating team + at least one other activity & requiring on-going negotiation/correspondence lasting >1 week (**Up to 8 hours in total**).

***Travel time:***

- None – onsite visit
- Up to 1 hour
- 1-2 hours
- 2-3 hours
- > 3 hours

After initial piloting in four centres with outreach services, 7 specialist rehabilitation services collected data on consecutive outreach assessments for a period of 3 months.

- Of a total of 731 assessments, 486 were assessments with a view to admission and 254 were for specialist community outreach advice/intervention.
- The mean age of the patients was 53 (sd 18) years; 547 (75%) had acquired brain injury, 70 (10%) had spinal cord injury, 62 (8%) had progressive neurological conditions and 26 (3%) had peripheral neurological conditions.
- A Consultant in Rehabilitation Medicine was present at over half of all referrals 407/731 (56%), and a more junior medic (trust grade or specialist registrar) was present at a further 20 (3%) assessments.

- There were 304/731(41%) assessments completed without any medical input (Consultant or non-Consultant).
  - 10 of these and were completed by multi-disciplinary teams (2-3 disciplines).
  - The remaining 294 assessments were uni-disciplinary:
    - 99 were specialist community outreach assessments, of which the majority (n=77) were completed by a senior Physiotherapist.
    - The remaining 195 assessments were for an admission assessment, completed by either by a senior nurse or senior allied health professional (AHP).

***Complexity and travel time***

Over three quarters (591/731- 81%) of the assessments were 'standard complexity' (up to 1 hour), with a further 125 (17%) reported as 'extended/up to 2 hours'. Thirteen assessments were reported as 'complex' and 2 'highly complex'.

A third of all assessments (242/731 – 33%) were completed "on site" – either the patient travelled to the unit or the team reviewed a patient on another ward/unit within the Trust. 489 assessments were completed off site, of which 289 (59%) were reached within 1 hour. For assessments that had a long travel time (greater than 3 hours) the time ranged from 4-8 hours with a median of 4 hours.

Table 7.13 shows the distribution of travel time for the seven outreach teams. The majority clearly operated mainly on a local basis rarely travelling for more than 1 hour, whilst 2 teams operated on a regional catchment and had much longer travel times.

***Distribution of travel time for the seven Outreach teams***

Table 100 shows the outcome from the outreach visit. Admission to the assessing unit was considered appropriate for 309/731 (42%) assessments; 181 of which were ready for admission at the time of assessment. Onward referral to a another Specialist Rehabilitation Service (Level 1/2) was recommended for 32 assessments (4%).

**Table 100: Outcome from assessment**

| Outcome | Total | % |
|---|---|---|
| Admit –ready for transfer | 181 | 25% |
| Admit – not yet ready | 128 | 18% |
| Review | 190 | 26% |
| Refer to community | 20 | 3% |
| Refer to Level 1/2a | 17 | 2% |
| Refer to Level 2b | 15 | 2% |
| Refer to Level 3 | 6 | 1% |
| Refer to slow stream | 7 | 1% |
| Refer Specialist Nursing home | 10 | 1% |
| Refer to other rehab | 6 | 1% |
| Refer to other specialty | 13 | 2% |
| Close | 117 | 16% |
| Unknown | 19 | 3% |

| Travel time | C023 | C029 | C031 | C054 | C064 | C077 | C109 | Total | % |
|---|---|---|---|---|---|---|---|---|---|
| No travel/on site | 35 | 68 | 15 | 1 | 19 | 99 | 5 | 242 | 33% |
| Up to 1 hour | 4 | 46 | 34 | 87 | 40 | 49 | 29 | 289 | 40% |
| 1-2 hours | | 10 | 34 | 5 | 2 | 5 | | 56 | 8% |
| 2-3 hours | | 21 | 78 | 5 | | 1 | | 105 | 14% |
| > 3 hours | | 21 | 15 | 2 | | 1 | | 39 | 5% |
| Total | 39 | 166 | 176 | 100 | 61 | 155 | 34 | 731 | 100% |

***Costs:***

Costs were calculated based on the 2013 unit costs (Curtis, 2013)[207] using an interactive spread sheet as illustrated in Figure 61. Tables 101, 102 and 103 show the cost matrices for assessment by a rehabilitation medicine consultant, a Trust grade doctor and a consultant grade AHP respectively.

| Tariff Calculation sheet | | | | | | |
|---|---|---|---|---|---|---|
| RCS-E Band | Cost Multiplier | Enter annual Bed days Here<br>bed days per annum | Weighted bed days | Tariff | % of total bed days | Total cost |
| Very High | 1.750 | 510 | 893 | £567.12 | 13% | £289,341 |
| High | 1.375 | 2183 | 3002 | £445.59 | 55% | £972,802 |
| Standard | 1.000 | 941 | 941 | £324.07 | 24% | £305,042 |
| Low | 0.700 | 225 | 158 | £226.85 | 6% | £51,139 |
| Very Low | 0.625 | 95 | 59 | £202.54 | 2% | £19,225 |
| | | 3955 | 5053 | | | **£1,637,550** |
| | £1,637,550 | **£414.05** | **£324.07** | | | |
| | **(Enter total annual** contract cost here) | **OBD cost** | **WBD Cost** | | | |

**Figure 61: Example of spreadsheet calculator**

**Table 101: Outreach costing matrix for assessment by a consultant only**

| Consultant only | | | | |
|---|---|---|---|---|
| | Costs based on £139 per hour (Curtis 2013) | | | |
| **Clinical activity** | Standard | Extended | Complex | Highly complex |
| Assessment | £139 | £278 | £556 | £1,112 |
| Clinical admin | £139 | £139 | £278 | £278 |
| Secretary | £25 | £25 | £38 | £50 |
| | £303 | £442 | £872 | £1,440 |
| | **Final tariff matrix inclusive of travel time** | | | |
| **Travel Time** | **Standard** | **Extended** | **Complex** | **Highly complex** |
| **1hr** | **£475** | **£614** | **£1,043** | **£1,612** |
| **2hrs** | **£626** | **£765** | **£1,195** | **£1,763** |
| **3 hrs** | **£778** | **£917** | **£1,346** | **£1,915** |
| **4 hrs** | **£929** | **£1,068** | **£1,498** | **£2,066** |

**Table 102: Outreach costing matrix for assessment by a Trust Grade Doctor**

| Associate specialist (Trust Grade doctor) only | | | | |
|---|---|---|---|---|
| | Costs based on £121 per hour (Curtis 2013) | | | |
| **Clinical activity** | Standard | Extended | Complex | Highly complex |
| Assessment | £121 | £242 | £484 | £968 |
| Clinical admin | £121 | £121 | £242 | £242 |
| Secretary | £25 | £25 | £38 | £50 |
| | £267 | £388 | £764 | £1,260 |
| | **Final tariff matrix inclusive of travel time** | | | |
| **Travel Time** | **Standard** | **Extended** | **Complex** | **Highly complex** |
| **1hr** | **£421** | **£542** | **£917** | **£1,414** |
| **2hrs** | **£554** | **£675** | **£1,051** | **£1,547** |
| **3 hrs** | **£688** | **£809** | **£1,184** | **£1,681** |
| **4 hrs** | **£821** | **£942** | **£1,318** | **£1,814** |

**Table 103: Outreach costing matrix for assessment by a consultant grade Allied Health Professional**

| Consultant Allied Health Professional (AHP) only | | | | |
|---|---|---|---|---|
| | Costs based on £78 per hour ( no costs in Curtis) | | | |
| **Clinical activity** | Standard | Extended | Complex | Highly complex |
| Assessment | £78 | £156 | £312 | £624 |
| Clinical admin | £78 | £78 | £156 | £156 |
| Secretary | £25 | £25 | £38 | £50 |
| | £181 | £259 | £506 | £830 |
| | **Final tariff matrix inclusive of travel time** | | | |
| **Travel Time** | **Standard** | **Extended** | **Complex** | **Highly complex** |
| **1hr** | **£292** | **£370** | **£616** | **£941** |
| **2hrs** | **£382** | **£460** | **£707** | **£1,031** |
| **3 hrs** | **£473** | **£551** | **£797** | **£1,122** |
| **4 hrs** | **£563** | **£641** | **£888** | **£1,212** |

**Table 104: Mean cost of outreach assessments across the different services**

| Unit | Range | Mean (sd) | Median (IQR) |
|---|---|---|---|
| **C023** | £267-£615 | £342 (£93) | £303 (£303) |
| **C029** | £181-£2160 | £524 (£351) | £459 (£303, £641) |
| **C031** | £181-£1773 | £614 (£315) | £615 (£373, £778) |
| **C054** | £181-£1643 | £532 (£220) | £449 (£459, £709) |
| **C064** | £181-£1340 | £406 (£229) | £303 (£292, £510) |
| **C077** | £181-£1598 | £256 (£178) | £181 (£181, £292) |
| **C109** | £267-£1238 | £700 (£261) | £701 (£528, £838) |

The cost of each assessment is dependent on the number of disciplines present, the assessment time and the travel costs, therefore costs vary, but the mean costs for each of the centres ranged from £256 to £700 per visit - see Table 104.

Table 105 illustrates the cost implications of the admissions avoided by outreach assessment. A total of 211 assessments did not lead to admission. At a mean cost of £482, the total cost of these 211 outreach assessments was £101,702. But if on average each of these avoided an unnecessary 10-day admission (in some cases a very conservative estimate) at an average cost of £5,300 the total savings in avoided admissions would be £1,075,900 – i.e. a total saving of nearly £1m.

**Table 105: Cost implications of admissions avoided**

| Service | No. of admissions avoided | Mean cost of assessment | Total cost of assessments for avoided admissions | Cost of a 4-6 week admission | Total cost of avoided admissions |
|---|---|---|---|---|---|
| **C023** | 8 | £342 | £2,736 | £17,500 | £140,000 |
| **C029** | 43 | £524 | £22,532 | £17,500 | £752,500 |
| **C031** | 88 | £614 | £54,032 | £17,500 | £1,540,000 |
| **C054** | 17 | £532 | £9,044 | £17,500 | £297,500 |
| **C064** | 20 | £406 | £8,120 | £17,500 | £350,000 |
| **C077** | 30 | £256 | £7,680 | £17,500 | £525,000 |
| **C109** | 4 | £700 | £2,800 | £17,500 | £70,000 |
| | **211** | **£482** | **£101,702** | **£17,500** | **£3,692,500** |

In summary: In the course of this exercise we tested a costing matrix for outreach services. The matrix was found to reflect the range of real life practice and published unit costs were used to price up the cost of staff time, allowing for different grades of staff.

Whilst at first sight this may seem an expensive service, its cost would readily be offset by the savings arising from even a small number of unnecessary admissions. These costings were shared with NHSE commissioners through the CRG for specialist rehabilitation in April 2015.

## *In summary for this Chapter*

In this part of the programme we developed patient level-costing protocols and applied these in different specialist rehabilitation settings to determine the differential treatment costs associated with different levels of caseload complexity in the UK.

Our key deliverables were:

1. Development and implementation of a novel weighted bed-day costing and payment model for specialist rehabilitation to derive indicative tariffs for specialist rehabilitation.
2. Implementation of the model to obtain accurate costing data to underpin a set of tariffs for specialist Level 1 and 2 rehabilitation services.
3. Evaluation of the cost impact of mandating the tariffs will have been provided to NHS England

These have all been achieved.

## *Changes from the original application*

1. We hoped to provide accurate costing data for the Level1a cognitive behavioural rehabilitation services. This has proved difficult, as NHS England currently commission only three of these services, with markedly different models.

   Broadly we found the Cognitive behavioural rehabilitation services separate into two types – post-acute hospital-based services and slower-stream community based services.

- The hospital-based services had significantly higher staff, as many of these patients require active neuropsychiatric interventions (e.g. management of psychotic symptoms, treatment under the Section of the Mental Health Act) and 1:1 supervision for challenging behaviours.
- The slower stream services tend to treat patients with less acute psychiatric needs, but a requirement for psychological surveillance / behavioural management over a longer period.

A further group of services provide highly specialist services for patients with very severe challenging behaviours / forensic psychiatry, but currently these tend to be commissioned under the Mental Health tariffs and do not report to UKROC

2. We also hoped to be able to provide a comprehensive evaluation of costs for slow-stream rehabilitation services in the community in the way that we have for the Level 1 and 2 services. In the event these have not been included in the NHS England service specification, so the UKROC data collection is not mandated. Therefore although the data have been assembled and shared, this area of community base rehabilitation provision is still lagging behind in tariff development as the data are not so complete or robust.

It is however an important area, and NHS England has increasingly become aware of the substantial on-going costs that this group of patients represent within the NHS continuing care services and the BSRM has introduced guidelines for different levels of nursing care

Within the life time of this programme, we have progressed this work as far as we are able, but there is undoubtedly a need for further detailed information about the exact client group attending these services, the therapies and interventions offered to meet their need, and the outcomes, using tools that are appropriate for the scale and type of changes observed. This will be the subject of a follow-on grant application.

# Chapter 8: National database development

## Objective:

To establish a nationwide database for centralised collation and analysis of case-episode data on needs, inputs, costs and person-centred outcomes from specialist neurorehabilitation services in the UK. Prospective data collection will inform tariff costs and provide ongoing benchmarking of quality, as well as evaluation of clinical benefits and cost-effectiveness.

## Deliverables

- All Level 1 and 2a services will be registered and routinely submitting high quality case-episode data to the UKROC dataset.
- Routine reporting and feedback systems established to provide contributing centres with activity analysis and benchmarking of outcomes.
- Final analysis of cost predictors and outcome in relation to complexity and other factors (e.g. dependency) will have been undertaken on the cleaned 2013/14 dataset.

## The UKROC database

The UK Rehabilitation Outcomes Collaborative was established in 2008 and the UKROC database established in 2009. The full UKROC dataset represents the inpatient rehabilitation subset of the Long Term neurological Conditions dataset, which was published by the NHS Information Centre in 2009. It comprises 30 items of demographic and process data for each admitted case episode together with measurement of:

1. **Rehabilitation needs** – documenting the individual requirements for rehabilitation using the Rehabilitation Complexity Scale (Currently the RCS-Ev13).
2. **Input** – documenting the nursing and therapy actually provided to meet those needs using the Northwick Park nursing and Therapy Dependency scales.
3. **Outcomes** – documenting the functional gains that are made during rehabilitation either using the Barthel Index,[54] the FIM and/or UK Functional Assessment Measure (UK FIM+FAM).[55]

As noted in *Chapter 3*, the dataset was designed to be hierarchical to minimise the burden of data collection. However, over 90 % of Level 1 and over 70 % of Level 2 services now collect and report the full UKROC dataset.

### *Software and training*

The UKROC software was developed in Microsoft Excel for the following reasons:

- The existing dedicated software for the Northwick Park Dependency Scores/Care Needs Assessment and the UKFIM+FAM that formed the basis of the software had already been programmed in Excel.
- The programme was widely available in most Trusts and so did not require the purchase / installation of additional software involving executable files or libraries.

The software sits on a computer/server within each Provider unit and includes patient identifiable fields for local use. When the data are exported for transfer to the UKROC database, they are pseudonymised so that the central UKROC database does not include patient-identifiable data. Name, NHS number, date of birth, address (-except first three letters of the postcode) are all excluded. Only the UKROC ID is retained for the purposes of re-identification by the treating service in case data require correction. They are sent to the UKROC team through the NHS-net or other suitably secure email link, and uploaded into the UKROC database which is securely held on a single server in Northwick Park Hospital, with full data protection in accordance NHS Information Policies. This arrangement was designed to address the concerns expressed by some Trusts about submitting patient data online.

The first edition of the UKROC software programme was developed during 2009 along with a detailed user guide. Training on data entry and use of the software was included the training courses for use of the UKROC tools. These courses were conducted in central workshops at Northwick Park throughout the programme, and also regionally in a series of road shows and workshops around England. Courses were free for staff from UKROC registered units. The Database was formally opened for use from April 2010.

## Uptake

Initially, data submission was optional, but the DH's Specialised Services National Definition Set third edition published in 2009 made it clear that all Level 1 specialist rehabilitations services would be required to report the full national clinical dataset, and Level 2 services at least the minimum dataset, for all in-patient episodes annually. Figure 62 shows the number of services in each class registered and reporting data to UKROC year by year.

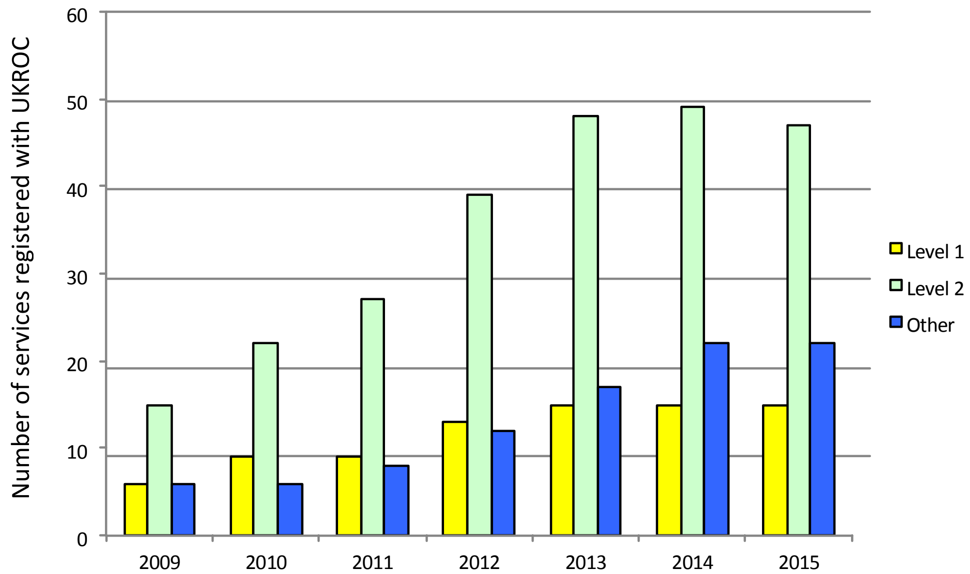



**Figure 62: The number of services registered and reporting data to UKROC by year**

- By 2011, nine Level 1 and 27 Level 2 services were registered and regularly submitting data. These included all eight Level 1 services in London.
- By 2013/14, when NHS England became responsible for commissioning specialised rehabilitation services, the service specification stipulated that all Level 1 and 2a services must be registered with UKROC and submitting the full UKROC dataset for each admitted episode in order to qualify for designation as a Level 1/2a service. Indeed only activity that was reported through UKROC would be counted for reimbursement. The Minimum dataset requirements are shown in *Table 106.*
- The PbR Guidance for 2013/14 mandated the multi-level weighted bed day currency and published indicative tariffs for Level 1 and 2 services. The same rules were carried forward in 2014/15.

## Table 106: UK ROC Minimum Data Reporting Requirements Checklist - 2013/14

| **Items** | **Service Level** (actual or aspired) | | | | | **Notes** |
|---|---|---|---|---|---|---|
| | **1*** | **2a*** | **2b*** | **2b** | **Other** | * using weighted bed day tariff |
| **Patient Identification & Demographics** | | | | | | |
| Patient Name | ✓ | ✓ | ✓ | ✓ | ✓ | for local use only |
| Date of Birth | ✓ | ✓ | ✓ | ✓ | ✓ | for age calculations + commissioners |
| Gender | ✓ | ✓ | ✓ | ✓ | ✓ | |
| Ethnicity | | | | | | desirable if available |
| Local Identifier | | | | | | for local use + commissioners only |
| Hospital Number | | | | | | for local use + commissioners only |
| NHS Number | ✓ | ✓ | ✓ | ✓ | | for local use + commissioners only |
| **Commissioning & Referral** | | | | | | |
| Funding Source (NHS CB, CCG, private etc) | ✓ | ✓ | ✓ | ✓ | ✓ | |
| Service Level (1, 2a, 2b, 3) | ✓ | ✓ | | | | if commissioned at several levels |
| Patient Category (a, b, c, d) | ✓ | ✓ | | | | |
| CCG name or code | ✓ | ✓ | ✓ | ✓ | ✓ | |
| GP Practice name, code and/or postcode | ? | ? | ? | ? | | may be required by commissioners |
| GP name and/or code | ? | ? | ? | ? | | may be required by commissioners |
| Patient postcode | | | | | | optional, though useful if available |
| Referral date | ✓ | ✓ | | | | |
| Referral source | ✓ | ✓ | | | | |
| Date of decision (added to active waiting list) | ✓ | ✓ | | | | |
| Date fit for admission | ✓ | ✓ | | | | |
| **Initial Assessment** | | | | | | |
| Date of initial assessment | ✓ | ✓ | | | | |
| Assessed by (uni/multi-disciplinary) | ✓ | ✓ | | | | |
| **Diagnosis** | | | | | | |
| Onset date (original and/or current) | ✓ | ✓ | | | | |
| Diagnosis category/subcategory | ✓ | ✓ | ✓ | ✓ | ✓ | |
| ICD 10 codes | | | | | | optional |
| **Admission Details** | | | | | | |
| Date of admission | ✓ | ✓ | ✓ | ✓ | ✓ | |
| Proposed discharge date | ✓ | ✓ | | | | |
| Proposed trimpoint date | | | | | | |
| Admitted from | ✓ | ✓ | | | | |
| Admission purpose | ✓ | ✓ | | | | |
| **Interruptions & Extensions** | | | | | | |
| Interruptions (start & end date, reason) | ✓ | ✓ | ✓ | ✓ | | |
| Extension date | ✓ | ✓ | ✓ | ✓ | | |
| **Discharge Details** | | | | | | |
| Date fit for discharge | ✓ | ✓ | | | | |
| Discharge date | ✓ | ✓ | ✓ | ✓ | ✓ | |
| Reason for delay | ✓ | ✓ | | | | |
| Discharge mode | ✓ | ✓ | | | | |
| Discharge destination | ✓ | ✓ | | | | |
| Discharge postcode | | | | | | optional, though useful if available |
| **Admission & Discharge Assessments** (all assessments should be submitted with fully itemised scores) | | | | | | |
| Patient Categorisation Tool (on admission) | ✓ | ✓ | | | | complexity measure |
| RCS-E (version 13 or 13T) – scored retrospectively | ✓ | ✓ | ✓ | ✓ | ✓ | complexity measure |
| FIM+FAM (including NIS) | ✓ | ✓ | | | | outcome measure |
| FIM, FIM+FAM, NPCNA or NPDS-H | | | ✓ | ✓ | | outcome measure |
| Barthel, FIM, FIM+FAM, NPCNA or NPDSH | | | | | ✓ | outcome measure |
| **Fortnightly Assessments** (scored retrospectively for all patients throughout the year based on what was provided) | | | | | | |
| RCS-E (version 13 or 13T) | ✓ | ✓ | ✓ | | | complexity/inputs measure |
| **Cross-Sectional Data Tranches** (all assessments should be scored retrospectively based on what was actually provided) Collected fortnightly for ALL patients until at least 100 sets of matching assessments have been completed | | | | | | |
| Matching RCS-E, NPCNA (or NPDS-H) & NPTDA | ✓ | ✓ | | | | complexity/inputs measures |
| RCS-E version 13 | | | ✓ | | ✓ | complexity/inputs measure |
| **Data Submission Frequency** | | | | | | |
| Monthly (including all current inpatients) | ✓ | ✓ | ✓ | | | |
| Quarterly (ideally including all current inpatients) | | | | ✓ | | monthly submissions preferred |

### *Recruitment against target*

The programme was registered with the Comprehensive Local Research Network (CLRN) and Figure 63 shows the quarter-by-quarter registered accruals.

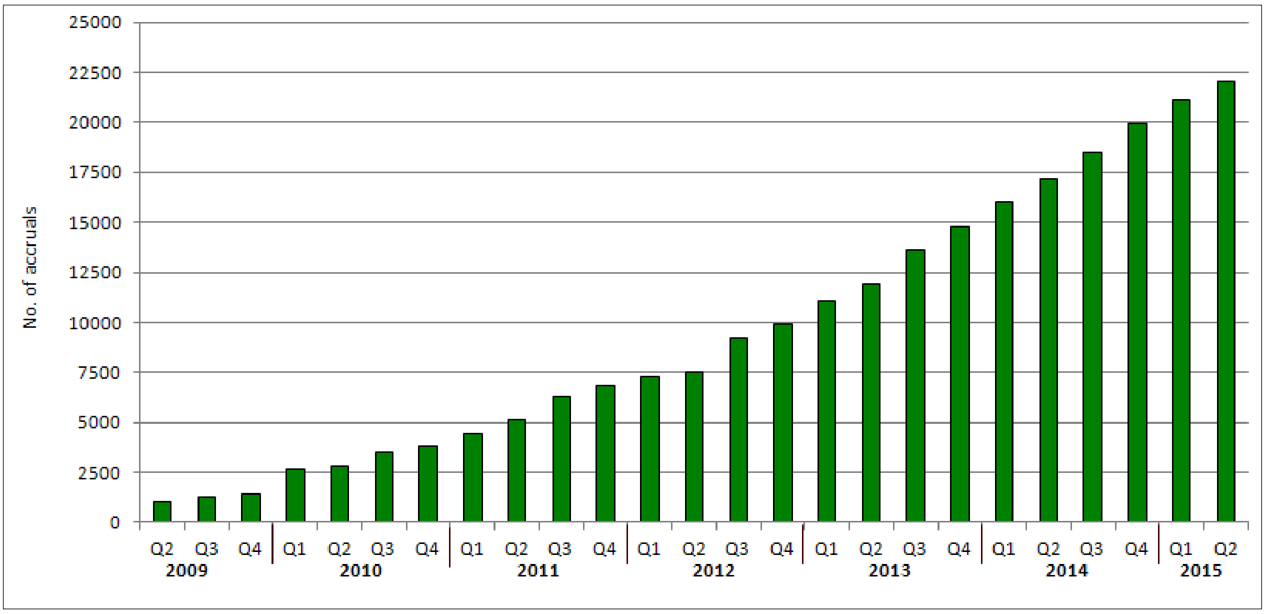



**Figure 63: Quarter-by-Quarter cumulative total accruals registered with the Comprehensive Local Research Network**

The original target for recruitment was n=12,500 by March 2014. That figure was exceeded by August 2013). However by that time the programme end date had been extended due to the radical restructuring of the NHS. The extended target was 22,000 by July 2015. Both the number of cases recruited, and the completeness of data collection have continued to grow year-on-year. The number of case episodes that met the minimum information requirements for accrual through the CLRN each year are shown in *Table107*.

**Table 107: Case episodes registered in the UKROC database since 2009**

| Month | Year | Accruals | Cumulative |
|---|---|---|---|
| Sept | 2009 | 1306 | 1306 |
| Sept | 2010 | 2205 | 3511 |
| Sept | 2011 | 2743 | 6254 |
| Sept | 2012 | 2945 | 9199 |
| Sept | 2013 | 4389 | 13588 |
| Sept | 2014 | 4887 | 18475 |
| May | 2015 | 3530 | 22005 |
| **Total** | | **22005** | |

### *Minimum reporting standards*

From July 2011, the UKROC database provided the commissioning dataset for the consortium of eight specialist rehabilitation services commissioned by the Pan-London Specialised Commissioning Group (SCG). From April 2012, when NHS England assumed responsibility for commissioning all level 1/2a services in England, the UKROC database provided the official commissioning dataset for all NHSE-commissioned services. Providers submit activity data each month, which is checked and feedback provided on any missing data. Providers then have one month to make any corrections to the data before the submission is frozen.

# Outcomes and Cost-efficient models of rehabilitation

Rehabilitation that helps an individual to improve their independence may be expected to reduce the long-term cost of providing care for them in the community. However, for this process to be 'cost-efficient', the initial investment in rehabilitation must be offset by the on-going savings on care within a reasonably short time-frame. Given increasing constraints on healthcare funding, providers are under mounting pressure to demonstrate the value for money of rehabilitation programmes and to report data on cost-efficiency as part of routine clinical practice. In 2006, we published a single centre study demonstrating the cost efficiency of rehabilitation and illustrating the differences between the UK and other countries when demonstrating this.[32]

As noted earlier in this report, the Functional Independence Measure (FIM) is widely used in the western world to demonstrate the functional gains achieved by individuals during rehabilitation. Studies in the 1990s equated points of FIM-gain with saved minutes of care,[208, 209] and thus 'FIM-efficiency' (calculated from 'FIM-gain from admission to discharge'/'length of stay') has been applied as a surrogate marker for cost-efficiency and used to benchmark the comparative efficiency of rehabilitation in different provider settings, and in different patient populations.[210-212]

In the UK, there has been less enthusiasm for this approach. This is partly because of concerns about the validity of mathematical manipulation of raw ordinal data, but also because of recognised floor and ceiling effects of the FIM™ [213, 214] which limit its use in some populations. For example, a very heavily dependent patient who progresses from needing help from two people for daily care, to needing only one person, may achieve a substantial reduction in the cost of continuing care, whilst changing very little on FIM rating. Similarly, an ambulant individual with severe cognitive deficits may achieve near-maximum FIM scores, but nevertheless need around-the-clock care to ensure their safety. So whilst global disability measures such as the FIM™ and Barthel Index (BI) are shown to correlate with care needs on a population basis,[49, 190] they cannot be used to assess them directly for a given individual.

The Northwick Park Dependency Score and Care Needs Assessment (NPCNA),[51, 76] on the other hand, have been specifically designed to measure care needs in the more dependent groups, and to provide a generic estimation of care hours and weekly cost of care in the community on an individual basis. If the cost of rehabilitation is known, and the savings in weekly cost of care estimated by the NPCNA, the time taken for savings to offset the cost of rehabilitation may offer a more direct indicator of cost-efficiency. This is particularly pertinent to the UK where the costs of life-long care in the community may be met by health and social services – especially in the more dependent cases.

In this section, we present a cohort analysis of the UKROC database examining functional outcome, change in care needs and cost-efficiency following specialist rehabilitation for patients with complex disability arising from neurological conditions. In particular we wished to determine whether the above findings of the single centre study of patients with acquired brain injury reported in 2006[32] are reproducible across multiple centres and across a wider range of neurological conditions. We therefore compare the parameters across different diagnoses and different levels of dependency on admission. We were also interested in the factors that predict the costs of rehabilitation, other than complexity and length of stay.

## *Methods*

The measures of interest were:

- Functional outcome – as measured by the NPDS and the UK FIM+FAM on admission and discharge. The UK FIM+FAM includes the FIM. The NPDS is translated by a computerised algorithm into the NPCNA, which estimates care hours/week and the approximate cost of providing for care needs in the community, regardless of who provides the care.
- The approximate cost of the in-patient rehabilitation episode - calculated per patient as bed-day cost multiplied by length of stay in days. The cost per bed-day is calculated on updated data from our previously published cost analysis.[131] The mean per diem costs for the different levels of service were: 1a: £540, 1b: £483, 1c: £634, 2a: 452, 2b: £418. (NB These figures differ slightly from those in Table 97 because they include the Market Forces Factor (MFF) and are thus representative of the actual cost to commissioners and providers).
- Cost efficiency is calculated as the time taken to offset the cost of rehabilitation by the resulting savings in the cost of on-going care in the community. This is calculated from 'Mean episode cost of rehabilitation' divided by 'mean reduction in weekly cost of care' from admission to discharge, as estimated by the NPCNA.
- For the purpose of comparison, 'FIM-Efficiency' is calculated as 'Gain in total FIM score from admission to discharge / Length of stay (week)'. We also calculated FIM+FAM efficiency in a similar manner.

### Data Extraction

De-identified data were extracted for all recorded in-patient episodes for adults >16 years discharged during the 4-year periods between 1.4.2010 and 31.3.14, if they had:

- A neurological condition recorded in the diagnostic category.
- A length of stay 7-400 days (plausible admissions for rehabilitation, excluding cases admitted for brief in-patient evaluation or long term management).
- Valid UK FIM+FAM and NPDS ratings completed both within 10 days of admission and within the last week before discharge.
- Data were collated in MS Excel and transferred to SPSS v22 for analysis.

### Data handling and analysis

Because data reporting was voluntary (at least in the earlier stages), missing data were expected. No data were imputed. Given the large size of the dataset and long-ordinal nature of the measures (i.e. many scoring categories), data were described and analysed using parametric statistics.

We accept that some of the data (e.g. length of stay, costs of rehabilitation) were very significantly skewed and so should normally be analysed using non-parametric statistics. In an exploratory analysis, we applied both parametric and non-parametric analytical techniques, as also found in several sections of Chapter 4. These gave very similar results leading to the same conclusions, although the parametric results were slightly more conservative. As some of the analyses (e.g. regression) make parametric assumptions in any event, for the sake of consistency we have presented the results of parametric analysis throughout. Boot strapping was used to minimise the effect of skewed data. However, the results of non-parametric analysis are also available on request. In the analysis presented below:

- Paired T tests were used to compare significant differences between admission and discharge and 95% confidence intervals were calculated using bootstrapping with samples of n=1000.
- One-way ANOVA tests and post hoc analysis with Bonferroni correction and bootstrapping with samples of n=1000 were used to compare intergroup differences.
- Pearson Correlations were used to examine relationships between variables.
- To identify the principal predictors of episode cost and length of stay, simple bivariate analyses (Pearson r) were first conducted to identify candidate variables. Those with a correlation of r=0.30 or more were entered into a stepwise linear regression model.

### *Results*

From a total of 13,855 registered episodes, 10,973 represented admissions for rehabilitation in a neurological condition. Of these 5,057 had both a valid NPDS and FIM+FAM in admission and discharge. See *Figure 64* for a flow chart of the data extraction process.

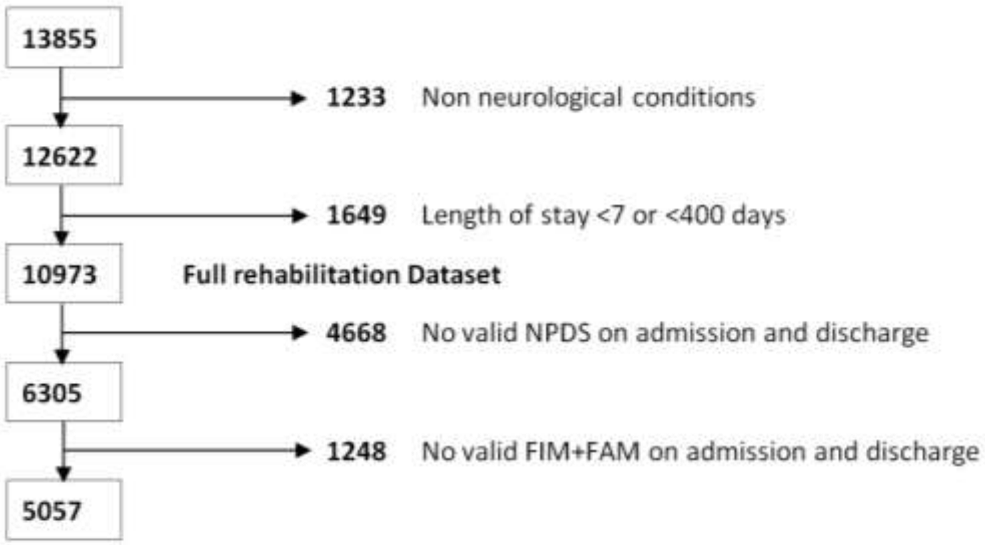



**Figure 64: Flow chart for data extraction**

# Demographics and diagnostic distribution

Demographics are given in *Table 108* for the whole sample and for the main diagnostic categories. Because the analysed sample contained less than 50% of the total rehabilitation dataset, demographics and key outcomes were first analysed for the total rehabilitation sample (n=10,093), and then for the sub-sample with complete NPDS and FIM+FAM scores (n=5057) – see *Table 109.*

Both samples comprised approximately 3:2 males:females, with a mean age at admission of 54.0 (sd=16.7) years, range 16-99 years. The mean rehabilitation length of stay was 77 days. Nearly three-quarters of the sample (71-72%) had acquired brain injury (ABI), the remained having spinal cord injuries (SCI) (10-11%), peripheral neurological conditions e.g. Guillain Barre Syndrome (6%) and progressive conditions e.g. multiple sclerosis (12%). The samples were also very similar in respect of their rehabilitation complexity, and their levels of functional independence and care needs, both on admission and discharge.

**Table 108: Demographics of the total and analysed populations**

| Parameter | Total rehabilitation dataset N=10,973* | | | | Analysis sample N=5057 | | | |
|---|---|---|---|---|---|---|---|---|
| **M:F ratio %** | 58/42% | | | | 59/41% | | | |
| | **Mean** | | **(SD)** | | **Mean** | | **SD** | |
| **Age (years)** | 54.4 | | (16.7) | | 54.0 | | (16.7) | |
| **Length of stay (days)** | 77.3 | | (62.6) | | 77.1 | | (60.1) | |
| **Cost of Episode (£)** | £35,217 | | (£30,936) | | £36,236 | | (£29,792) | |
| **Diagnostic category** | **N** | | **%** | | **N** | | **%** | |
| Acquired brain injury | 7899 | | 72% | | 3592 | | 71% | |
| Spinal cord injury | 1115 | | 10% | | 586 | | 11% | |
| Peripheral neurological | 614 | | 6% | | 229 | | 6% | |
| Progressive condition | 1345 | | 12% | | 580 | | 12% | |
| **Aetiology** | **N** | | **%** | | **N** | | **%** | |
| Trauma | 2173 | | 19.8% | | 1110 | | 21.9% | |
| Vascular | 4349 | | 39.6% | | 1852 | | 36.6% | |
| Inflammatory | 887 | | 8.1% | | 443 | | 8.8% | |
| Tumour | 673 | | 6.1% | | 339 | | 6.7% | |
| Other | 1631 | | 14.9% | | 769 | | 15.2% | |
| Multiple sclerosis | 1023 | | 9.3% | | 459 | | 9.1% | |
| Motor neurone disease | 25 | | 0.2% | | 11 | | 0.2% | |
| Parkinson's disease | 65 | | 0.6% | | 28 | | 0.6% | |
| Missing | 147 | | 1.3% | | 46 | | 0.9% | |
| | **Admission** | | **Discharge** | | **Admission** | | **Discharge** | |
| | **Mean** | **(SD)** | **Mean** | **(SD)** | **Mean** | **(SD)** | **Mean** | **(SD)** |
| **Rehabilitation complexity** | **N=10,058** | | **N=9,742** | | **N=5010** | | **N=4987** | |
| RCS version 12 | 11.6 | (2.6) | 8.6 | (3.5) | 11.8 | (2.5) | 9.0 | (3.5) |
| **Barthel Index** | **N=10,159** | | **N=9,784** | | **N=5057** | | **N=5057** | |
| BI total | 8.0 | (6.1) | 12.2 | (6.8) | 7.4 | (6.0) | 11.7 | (6.8) |
| **FIM+FAM** | **N=8356** | | **N=7902** | | **N=5057** | | **N=5057** | |
| Motor subscale | 54.2 | (28.6) | 75.9 | (31.1) | 52.6 | (28.0) | 74.6 | (31.1) |
| Cognitive Subscale | 61.2 | (25.2) | 73.2 | (23.1) | 60.7 | (25.4) | 72.6 | (23.4) |
| Total | 115.2 | (46.7) | 149.1 | (49.6) | 113.2 | (46.2) | 147.2 | (50.0) |
| **NPDS/NPCNA** | **N=6652** | | **N=6591** | | **N=5057** | | **N=5057** | |
| NPDS total score | 28.6 | (17.5) | 19.1 | (17.4) | 29.2 | (17.4) | 19.0 | (17.5) |
| Estimated care hrs/wk | 42.5 | (20.6) | 29.7 | (21.9) | 43.1 | (20.1) | 29.6 | (21.8) |
| Estimated care costs/wk | £1475 | (949) | £1000 | (954) | £1503 | (936) | £998 | (948) |

As these samples were similar in all respects, and the sub-sample with complete data was still large (representing >5,000 episodes), we used this dataset for the rest of our analysis. This way we could be reasonably confident that the analysis population was representative of the total rehabilitation group, but also that any reported between group differences were not simply due to sampling differences, reflecting the fact that different measures were used in different people.

### Functional gain and change in dependency

*Table 109* shows the change from admission to discharge for functional independence and dependency. There was a highly significant increase in all parameters of functional independence (FIM+FAM, p<0.001), with corresponding reduction in all parameters of dependency (NPDS/NPCNA, p<0.001). Individually calculated, the mean FIM efficiency (FIM gain/length of stay in weeks) was 3.1 (SD 3.8), the FIM+FAM efficiency was 4.8 (SD 6.0). The mean total cost of the rehabilitation programme was £36,236 (SD 27,792) and mean savings in ongoing cost of care in the community was £505 (SD 287) /week. The mean time taken to offset the initial costs of rehabilitation was therefore 18.7 (SD 137) months.

**Table 109: Dependency and functional outcome scores on admission and discharge (n=5057)**

| | Admission Mean (SD) | Discharge Mean (SD) | Mean difference | 95% CIs Upper | 95% CIs Lower | t | P value* |
|---|---|---|---|---|---|---|---|
| **Functional Independence (UK Functional Assessment Measure - FIM+FAM)** | | | | | | | |
| **Self-care** | 27.9 (12.8) | 35.5 (13.1) | 8.3 | 8.0 | 8.2 | 67.6 | <0.001 |
| **Sphincter** | 7.4 (4.8) | 9.9 (4.7) | 2.4 | 2.3 | 2.5 | 46.9 | <0.001 |
| **Transfers** | 11.3 (8.2) | 18.0 (9.0) | 6.7 | 6.5 | 6.9 | 68.4 | <0.001 |
| **Locomotion** | 6.6 (4.7) | 11.3 (6.0) | 4.6 | 4.5 | 4.8 | 68.3 | <0.001 |
| **Communication** | 23.0 (9.9) | 26.9 (8.9) | 3.9 | 3.8 | 4.1 | 50.1 | <0.001 |
| **Psychosocial** | 17.2 (7.2) | 20.6 (6.7) | 3.4 | 3.3 | 3.5 | 49.2 | <0.001 |
| **Cognition** | 20.5 (10.4) | 25.2 (9.4) | 4.6 | 4.5 | 4.8 | 52.2 | <0.001 |
| **Total Motor** | 52.6 (28.0) | 74.6 (31.1) | 22.0 | 21.5 | 22.6 | 75.5 | <0.001 |
| **Total Cognitive** | 60.7 (25.4) | 72.6 (23.4) | 12.0 | 11.6 | 12.4 | 58.5 | <0.001 |
| **Total FIM+FAM** | 113.2 (46.2) | 147.2 (50.0) | 34.0 | 33.1 | 34.8 | 78.1 | <0.001 |
| **FIM-Motor** | 43.1(24.4) | 61.2 (26.4) | 18.1 | 17.6 | 18.6 | 72.2 | <0.001 |
| **FIM-Cognitive** | 22.5 (9.7) | 26.4 (8.7) | 3.9 | 3.7 | 4.0 | 49.8 | <0.001 |
| **Total FIM** | 65.7 (30.0) | 87.6 (32.4 | 21.9 | 21.3 | 22.5 | 74.9 | <0.001 |
| **Dependency (Northwick Park Dependency Score and Care Needs Assessment – NPDS/NPCNA)** | | | | | | | |
| **Total NPDS** | 29.2 (17.4) | 19.0 (17.5) | -10.2 | -10.5 | -9.8 | -56.6 | <0.001 |
| **Care hrs/wk** | 43.1 (20.1) | 29.6 (21.8) | -13.5 | -14.0 | -13.1 | -57.6 | <0.001 |
| **Care costs/week** | £1503 (936) | £998 (948) | -£505 | -£528 | -£482 | -43.7 | <0.001 |

**two-tailed significance*

## *Analysis by diagnostic category*

*Table 110* shows the demographics of the analysis sample split by diagnostic category and aetiology.

**Table 110: Demographics of the analysis population (n=5057) by diagnostic category and aetiology**

| Parameter | ABI N=3592 (71%) | SCI N=586 (11%) | Peripheral N=299 (6%) | Progressive N=580 (12%) | All N=5057 |
|---|---|---|---|---|---|
| **Age Mean (SD)** | 53.0 (16.9) | 59.3 (16.6) | 54.4 (16.3) | 52.6 (13.8) | 54.0 (16.7) |
| **M:F ratio %** | 62/38% | 60/40% | 57/43% | 43/57% | 59/41% |
| **Length of stay (days) Mean (SD) days** | 83.0 (62.7) | 66.7 (50.0) | 69.7 (55.8) | 54.3 (46.5) | 77.1 (60.1) |
| **Cost of episode Mean (SD)** | £39,562 (£31,468) | £30,289 (£23,047) | £31,803 (£26,725) | £23,924 (£20,926) | £36,236 (£29,792) |
| **Diagnostic sub-categories - N(%)** | | | | | |
| **Trauma** | 970 (27.0%) | 134 (22.9%) | 6 (2.0%) | | 1110 (21.9%) |
| **Vascular** | 1774 (49.4%) | 61 (10.4%) | 17 (5.7%) | | 1852 (36.6%) |
| **Inflammatory** | 135 (3.8%) | 118 (20.1%) | 190 (63.5%) | | 443 (8.8%) |
| **Tumour** | 239 (6.7%) | 100 (17.1%) | - | | 339 (6.7%) |
| **Other** | 922 (12.4%) | 165 (28.2%) | 83 (27.8%) | | 769 (15.2%) |
| **Multiple sclerosis** | | | | 459 (79.0%) | 459 (9.1%) |
| **Motor neurone disease** | | | | 11 (1.9%) | 11 (0.2%) |
| **Parkinson's disease** | | | | 28 (4.8%) | 28 (0.6%) |
| **Missing** | | | | | 46 (0.9%) |

One-way Anova tests confirmed significant between group differences in age, length of stay and episode costs (p<0.001). Post hoc tests showed that SCI patients were significantly older than all other diagnostic categories (p<0.001), whilst patients with acquired brain injury had significantly longer lengths of stay and overall episode costs (p<0.001).

The differences in functional outcome between the different diagnostic groups are summarised in *Table 111.* On admission, there were no significant differences between any of the groups for FIM+FAM Motor score. As expected, post hoc tests showed FIM+FAM cognitive scores both on admission and discharge to be significantly lower in ABI than any of the other groups, (p<0.001). They were also significantly lower for Progressive conditions than for the SCI and Peripheral Neurology groups, but the latter were similar.

Between admission and discharge, change in FIM+FAM Motor score was sigificantly different between all groups (p<0.001), except between ABI and SCI (p=1.0). Change in FIM+FAM cognitive score was significantly different between all of the groups (p<0.01) except for SCI and Progressive conditions (p=1.0). Mean individually calculated FIM+FAM efficiency was highest the Peripheral neurology group and lowest in Progressive conditions. Post hoc tests confirmed significant differences between all groups (p≤0.01) except ABI and Peripheral conditions. Changes at item level for the four groups are illustrated in *Figure 65.*

**Table 111: Comparison of functional and dependency scores between diagnostic groups**

| Parameter | ABI n=3591 | | SCI n=586 | | Peripheral n=299 | | Progressive n=580 | | One-way ANOVA Between Groups | |
|---|---|---|---|---|---|---|---|---|---|---|
| **UK FIM+FAM** | | | | | | | | | | |
| **Admission** | **Mean** | **SD** | **Mean** | **SD** | **Mean** | **SD** | **Mean** | **SD** | **F** | **P** |
| Motor | 52.3 | (30.0) | 55.1 | (20.9) | 53.7 | (22.6) | 55.0 | (20.9) | 2.3 | 0.071 |
| Cognitive | 52.5 | (23.6) | 85.7 | (13.1) | 83.3 | (13.6) | 85.7 | (13.1) | 616.1 | <0.001 |
| Total | 104.7 | (48.1) | 140.8 | (28.6) | 137.0 | (31.1) | 140.8 | (28.6) | 164.2 | <0.001 |
| **Discharge** | | | | | | | | | | |
| Motor | 75.0 | (32.6) | 78.1 | (24.4) | 85.6 | (24.5) | 78.1 | (24.4) | 46.2 | <0.001 |
| Cognitive | 67.4 | (24.0) | 89.9 | (11.5) | 90.4 | (11.3) | 89.9 | (11.5) | 275.1 | <0.001 |
| Total | 142.4 | (53.2) | 168.0 | (31.8) | 176.0 | (32.5) | 168.0 | (31.8) | 85.0 | <0.001 |
| **Change** | | | | | | | | | | |
| Motor | 22.8 | (21.3) | 23.0 | (18.1) | 31.9 | (21.8) | 23.1 | (18.1) | 80.8 | <0.001 |
| Cognitive | 14.9 | (15.2) | 4.2 | (8.1) | 7.1 | (10.1) | 4.2 | (8.0) | 189.2 | <0.001 |
| Total | 37.7 | (32.7) | 27.2 | (21.7) | 39.0 | (27.0) | 27.4 | (21.7) | 104.6 | <0.001 |
| **FIM efficiency/week** | 3.2 | (3.9) | 3.0 | (3.5) | 4.1 | (3.8) | 2.0 | (3.3) | 21.6 | <0.001 |
| **FIM+FAM efficiency** | 5.2 | (6.3) | 4.2 | (4.6) | 5.8 | (5.6) | 3.0 | (4.3) | 27.4 | <0.001 |
| **NPDS/NPCNA** | | | | | | | | | | |
| **Admission** | **Mean** | **SD** | **Mean** | **SD** | **Mean** | **SD** | **Mean** | **SD** | **F** | **P** |
| NPDS | 30.5 | (18.2) | 25.0 | (13.8) | 25.3 | (14.4) | 26.8 | (15.2) | 27.9 | <0.001 |
| Care hours/wk | 43.5 | (20.3) | 41.7 | (19.1) | 41.7 | (19.5) | 43.9 | (20.7) | 2.0 | 0.114 |
| Care costs | £1,582 | (£938) | £1,314 | (£887) | £1,287 | (£909) | £1,391 | (£941) | 24.8 | <0.001 |
| **Discharge** | | | | | | | | | | |
| NPDS | 19.1 | (18.5) | 14.8 | (13.1) | 11.9 | (13.0) | 21.4 | (15.4) | 35.2 | <0.001 |
| Care hours/wk | 30.0 | (22.0) | 25.9 | (19.8) | 20.2 | (18.8) | 35.9 | (22.1) | 41.3 | <0.001 |
| Care costs | £1,059 | (£980) | £771 | (£799) | £606 | (£745) | £1,088 | (£917) | 35.3 | <0.001 |
| **Change** | | | | | | | | | | |
| NPDS | -10.6 | (13.1) | -10.2 | (11.2) | -13.3 | (13.3) | -5.4 | (9.7) | 35.9 | <0.001 |
| Care hours/wk | -13.4 | (16.3) | -15.7 | (16.1) | -21.5 | (18.5) | -8.0 | (14.1) | 50.4 | <0.001 |
| Care costs | -£518 | (£809) | -£539 | (£794) | -£671 | (£874) | -£305 | (£748) | 16.7 | <0.001 |
| **Time to offset costs of rehabilitation** | 20.3 | (156.6) | 17.7 | (83.7) | 17.5 | (53.8) | 11.5 | (77.8) | 0.56 | 0.640 |

*UK FIM+FAM = Functional Assessment Measure, NPDS = Northwick Park Dependency Score, NPCNA = Northwick Park Care Needs Assessment*

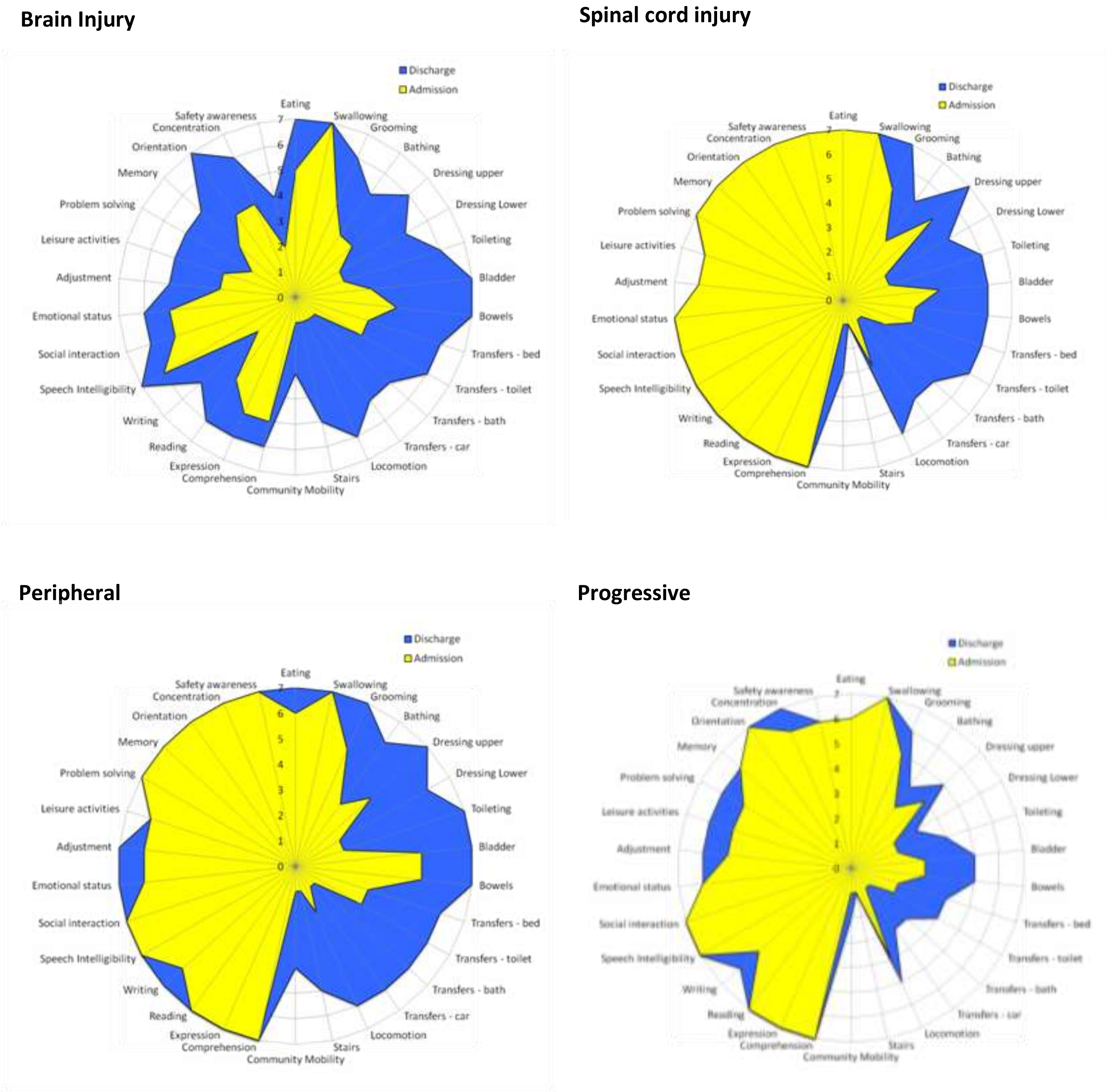



**Figure 65: FAM splats for the four diagnostic groups**

The differences in dependency are also summarised in *Table 112*. On admission, there were no significant differences between any of the groups for the estimated care hours per week. Post hoc tests showed NPDS scores and estimated weekly care costs to be significantly higher in ABI than any of the other groups ($p<0.001$) but there were no significant differences between any of the other groups ($p=1.0$).

Between admission and discharge reduction in dependency and care costs were sigificantly different between all groups ($p<0.05$), except between ABI and SCI ($p=1.0$). The mean individually calculated time to offset the cost of rehabilitation was lowest in the progressive conditions group and highest in ABI, but the data were widely spread and ANOVA did not show any significant between group differences ($p=0.64$).

## *Differences between groups based on dependency at admission*

The change in dependency, care needs and cost of care in the community are summarised in Table 112. grouped by the level of dependency on admission.

**Table 112: Comparison of costs and efficiency between dependency groups (n=5057)**

| Parameter | Low dependency N=767 (15%) | | Medium dependency N=1428 (28%) | | High dependency N=2862 (57%) | | One-way ANOVA | |
|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD | F | P |
| Length of stay | 47 | (40) | 59 | (47) | 94 | (65) | 309.1 | <0.001 |
| Cost of Rehabilitation | £23,075 | (£22,325) | £27,528 | (£22,994) | £44,107 | (£32,017) | 259.9 | <0.001 |
| **Admission** | | | | | | | | |
| NPDS | 4.67 | (3.1) | 17.1 | (4.3) | 41.8 | (11.3) | 7090.8 | <0.001 |
| Care hours/wk | 14.1 | (8.2) | 31.5 | (9.7) | 57.3 | (12.6) | 5506.5 | <0.001 |
| Care costs £/wk | £360 | (£369) | £935 | (£517) | £2,212 | (£700) | 3361.1 | <0.001 |
| **Discharge** | | | | | | | | |
| NPDS | 3.4 | (4.9) | 9.3 | (8.0) | 28.1 | (17.5) | 1423.3 | <0.001 |
| Care hours/wk | 8.8 | (8.3) | 18.6 | (13.8) | 41.1 | (2.8) | 1430.2 | <0.001 |
| Care costs £/wk | £238 | (£367) | £533 | (£560) | £1454 | (£969) | 1036.2 | <0.001 |
| **Change** | | | | | | | | |
| NPDS | -1.3 | (4.8) | 7.8 | (7.7) | 13.7 | (14.6) | 370.3 | <0.001 |
| Care hours/wk | -5.3 | (9.0) | 12.9 | (12.3) | -16.1 | (18.9) | 137.7 | <0.001 |
| Care costs £/wk | -£121 | (£428) | £402 | (£640) | -£665 | (£917) | 160.2 | <0.001 |
| **Efficiency** | | | | | | | | |
| Time to offset costs of rehab (months) | 38.4 | (157) | 19.7 | (119) | 13.6 | (140) | 7.7 | <0.001 |
| FIM Efficiency | 2.8 | (3.5) | 3.9 | (4.1) | 2.7 | (3.7) | 50.8 | <0.001 |
| FAM efficiency | 5.2 | (6.0) | 6.1 | (6.2) | 4.1 | (5.7) | 54.4 | <0.001 |

*UK FIM+FAM = Functional Assessment Measure, NPDS = Northwick Park Dependency Score, NPCNA = Northwick Park Care Needs Assessment*

As anticipated, length of stay and the total cost of admission were greatest in the high dependency group and smallest in the low dependency group, with significant differences seen between all groups on post hoc tests (p<0.001).

The ongoing care hours and costs of care in the community remained high at discharge in the same pattern as on admission, but the reduction in care hours and costs was greater in the higher dependency groups, reflecting the higher starting levels – again with significant differences between all dependency groups (p<0.001).

Depsite the higher cost of the rehabilitation, the time to offset the costs of treatment through savings in the cost of ongoing community care was shortest in the high, followed by the medium dependency group. Post hoc tests revealed that the difference reached significance between the medium and low dependency groups (p=0.02) but not between the high and medium groups.

By contrast, FIM efficiency was highest in the medium dependency group but similar between the low and high dependency groups (p=1.0). FIM+FAM efficiency was highest in the medum dependency group and lowest in the high dependency group, with significant differences seen between all three groups (p≤0.002).

## *In conclusion*

The findings from this multicentre analysis provide strong evidence for the overall cost-effectiveness of rehabilitation, and also confirm the generalisability of the conclusions from our previous single centre analysis.[32] Overall across the sample, the initial costs of rehabilitation were offset by savings in the on-going costs of care within just 18.7 months - which is comparatively short in this young adult population (mean age 54) who may be expected to survive on average for at least another 20-25 years.

Also, as in the previous study, the cost savings were most marked in the highly dependent patients, with the time to offset the costs just over 1 year (13.6 months). This does, of course, assume that the gains remain at least stable. Data captured during our related community-based follow-up study[215] and presented in Chapter 4 confirms that NPDS gains were maintained (and in some cases even improved further) for at least one year after discharge from a specialist Level 1 rehabilitation unit within this cohort.

This finding is important as many of these patients would not be considered eligible for rehabilitation in health systems that rely on the demonstration FIM efficiency, as the latter was higher in the medium dependency group. That said, because of the longer lengths of stay in this population of patients with complex disabilities requiring tertiary specialist rehabilitation, FIM efficiency was considerably lower than the normal for rehabilitation services in the US/Australia – which mainly comprise the equivalent of Level 2b/3 services in the UK (as discussed in Chapter 6).

This study also demonstrates that rehabilitation is cost-efficient not only for patients with acquired brain injury (as in the previously published study),[32] but for other neurological conditions as well. Perhaps an unexpected finding from this study was that the time to offset the costs of rehabilitation was shortest in the group of patients with progressive conditions (approximately 11 months compared to 18-20 months in the other diagnostic groups). Patients with progressive condition are often considered to have less rehabilitation potential, and indeed the cost savings (on average £305) were lower than for the other conditions. However, the cost of inpatient rehabilitation was also lower, due to the shorter lengths of stay and the fact that many of these patients were admitted to their local (Level 2) services. Overall, therefore, this investment proved financially to be very worthwhile, despite the poorer prognosis in the very long term. This is an important message for commissioners who are often reluctant to pay for such services.

# Cost-efficient models of in-patient rehabilitation – service levels

We also considered the outcomes and cost-efficiency across the different levels of service as shown in *Table 113* and *Figures 66 and 67 a-d*. *Figure 66* shows the FAM splats for the five different levels of service. Levels 2a and 2b show broadly similar profiles, which are not dissimilar from 1b services. Levels1 a and 1c show strikingly different profiles reflecting their polar differences in physical ability.

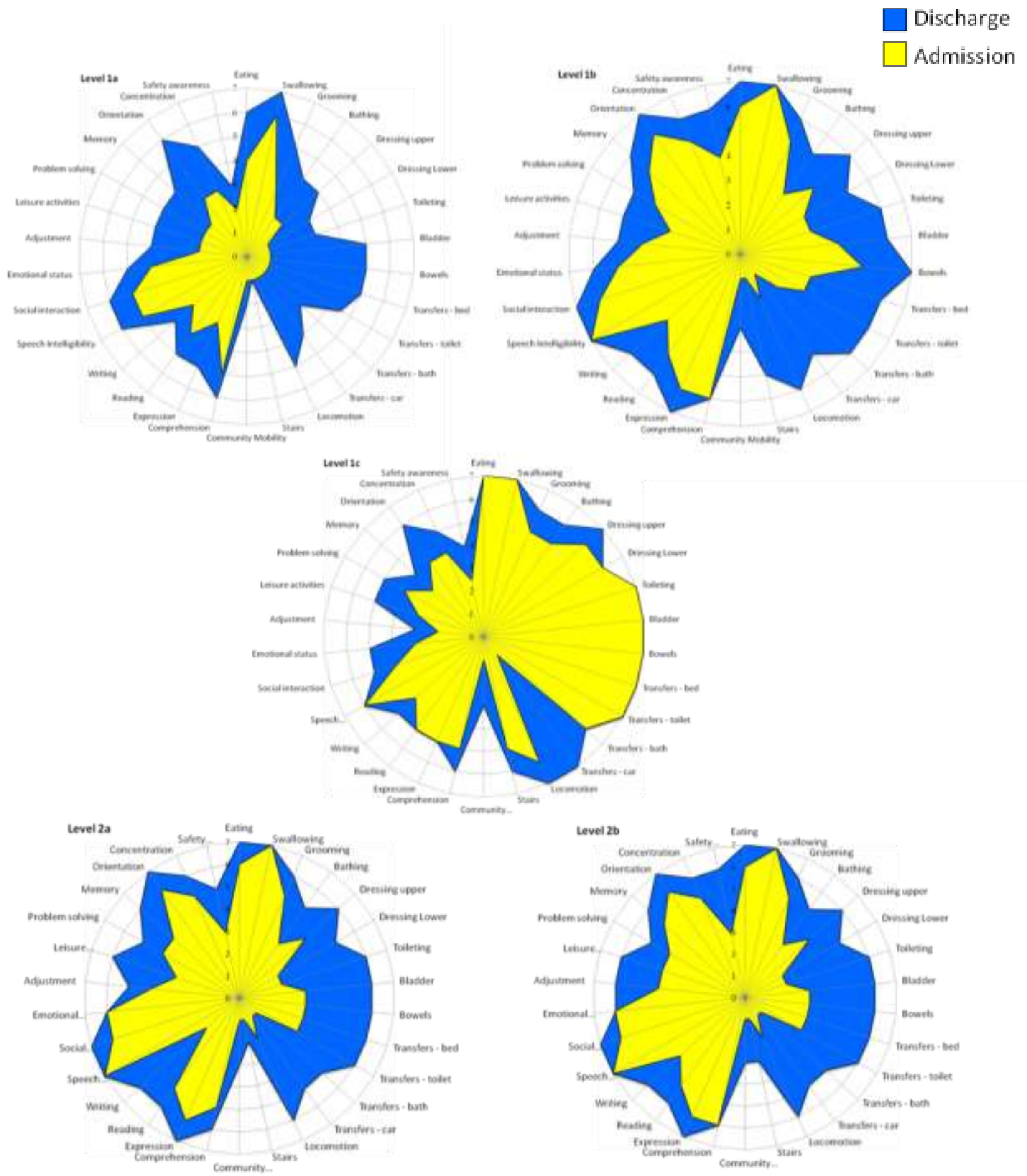



**Figure 66: FAM Splats for the five different service levels**

The episode costs mirrored length of stay fairly closely as shown in *Figure 67*. Level 1c (cognitive behavioural) rehabilitation services were clearly different from the other Level 1 and 2 services in respect of length of stay and therefore episode costs. In addition, the gains in FIM+FAM score and costs of care were substantially less (see *Table 115*) than for other services. This is not surprising (see discussion for this section, page 234) so the Level 1c services were excluded from the ANOVAs described below. We were nevertheless interested in the differences between the 1a, 1b, 2a and 2b services.

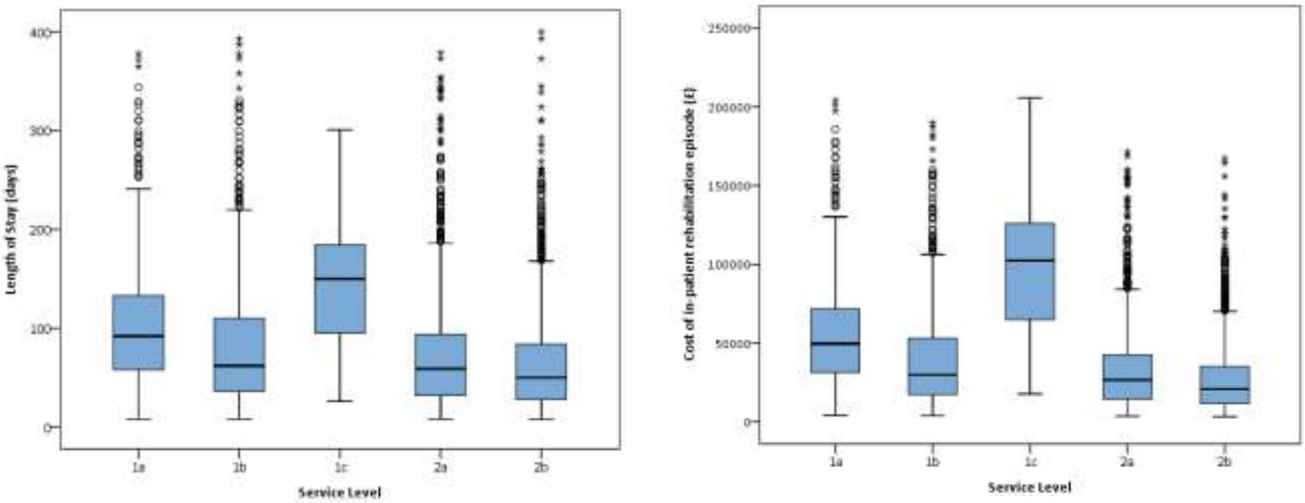
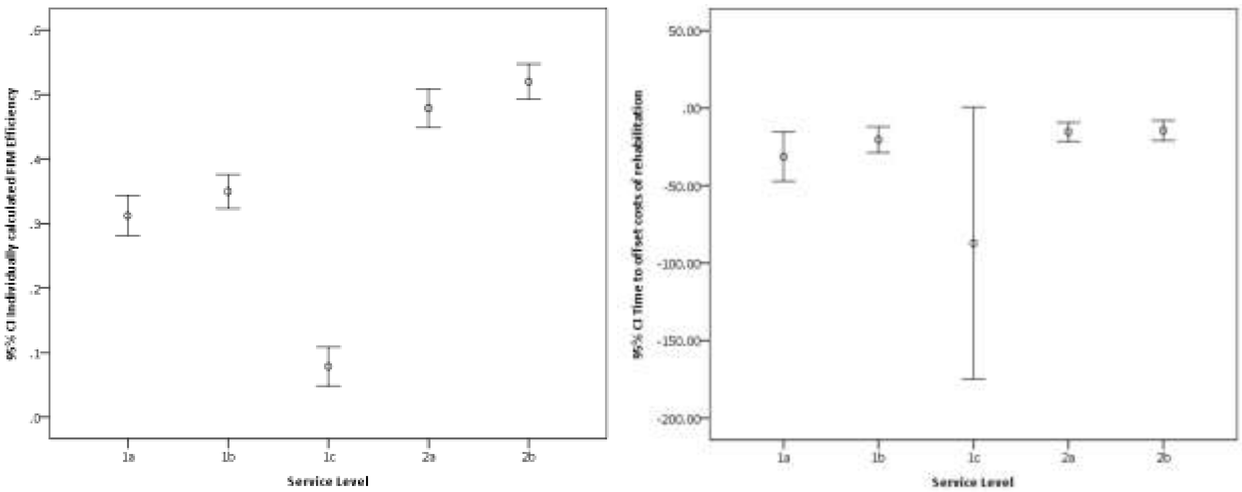



**Figure 67: Length of stay, episode costs, FIM efficiency and time to offset the cost of rehabilitation across the five different levels of service. Box plots show median, interquartile range and range.**

*NB: Please note that the time to offset costs is given in a negative value on these graphs, indicating the reduction in on-going costs of care. Higher (less negative) values indicate a shorter time to offset costs.*



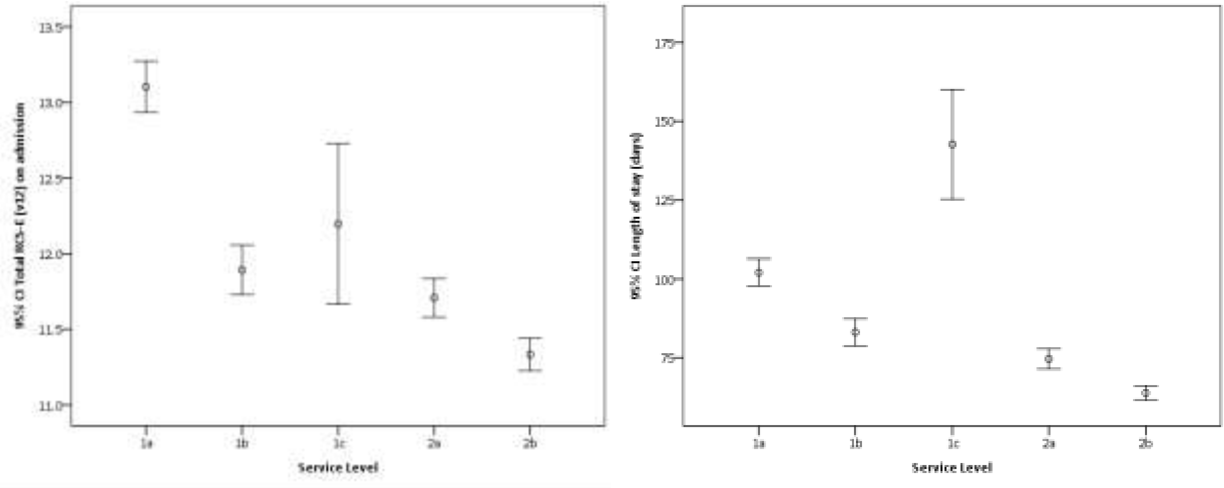



**Figure 68:** Mean and 95% Confidence intervals for Rehabilitation Complexity on admission and length of stay

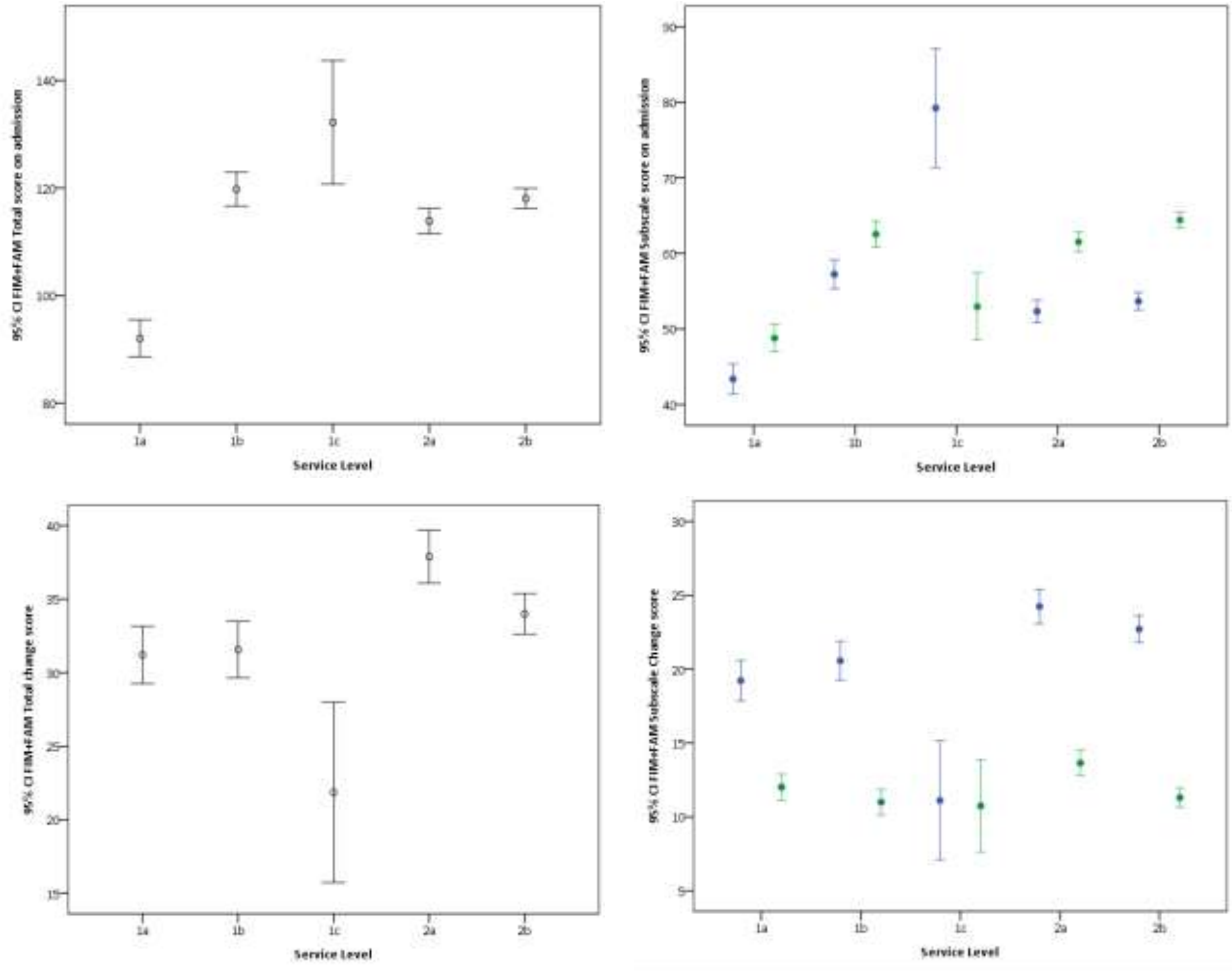



**Figure 69:** Mean and 95% Confidence intervals for FIM+FAM on admission and change scores

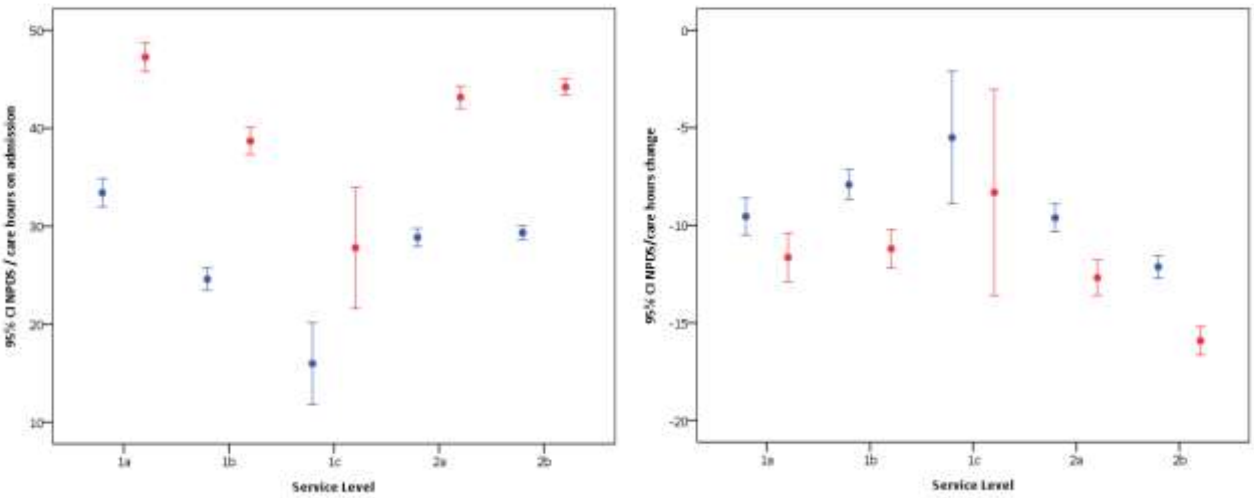



**Figure 70:** **Mean and 95% Confidence intervals for NPDS and care hours on admission and change scores**

**Table 113:** Comparison of functional independence and efficiency between different service levels

| Service level | 1a n=817 | | 1b n=902 | | 1c n=52 | | 2a n=1304 | | 2b n=1981 | | One-way ANOVA Between Groups (Excluding Level 1c) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD | F | P |
| **RCS-E v 12** | | | | | | | | | | | | |
| **Admission** | 13.1 | (2.4) | 11.9 | (2.5) | 12.2 | (1.9) | 11.7 | (2.3) | 11.3 | (2.4) | 78.1 | <0.001 |
| **Change** | -2.2 | (3.0) | -2.0 | (2.7) | -2.7 | (3.0) | -3.3 | (3.1) | -3.3 | (3.2) | 44.5 | |
| **UK FIM+FAM** | | | | | | | | | | | | |
| **Admission** | | | | | | | | | | | | |
| **Motor** | 43.4 | (28.9) | 57.2 | (29.3) | 79.3 | (28.2) | 52.3 | (27.0) | 53.7 | (26.4) | 40.0 | <0.001 |
| **Cognitive** | 48.7 | (26.2) | 62.5 | (26.3) | 53.0 | (16.0) | 61.5 | (24.0) | 64.4 | (24.1) | 80.0 | <0.001 |
| **Total** | 92.0 | (50.2) | 119.8 | (48.5) | 132.2 | (41.2) | 113.9 | (43.4) | 118.0 | (42.9) | 73.5 | <0.001 |
| **Discharge** | | | | | | | | | | | | |
| **Motor** | 62.5 | (35.4) | 77.8 | (31.1) | 90.4 | (26.7) | 76.6 | (29.2) | 76.4 | (29.4) | 49.6 | <0.001 |
| **Cognitive** | 60.8 | (27.2) | 73.5 | (24.3) | 63.7 | (17.9) | 75.2 | (21.1) | 75.7 | (21.3) | 91.3 | <0.001 |
| **Total** | 123.2 | (59.6) | 151.4 | (50.8) | 154.1 | (41.6) | 151.8 | (44.9) | 152.0 | (45.8) | 77.5 | <0.001 |
| **Change** | | | | | | | | | | | | |
| **Motor** | 19.2 | (19.8) | 20.6 | (20.1) | 11.1 | (14.5) | 24.2 | (21.6) | 22.7 | (20.8) | 12.1 | <0.001 |
| **Cognitive** | 12.0 | (12.8) | 11.0 | (13.3) | 10.8 | (11.2) | 13.7 | (15.6) | 11.3 | (15.0) | 8.6 | <0.001 |
| **Total** | 31.2 | (28.5) | 31.6 | (29.4) | 21.9 | (22.0) | 37.9 | (33.1) | 34.0 | (31.0) | 10.9 | <0.001 |
| **FIM efficiency** | 2.2 | (3.2) | 2.4 | (3.8) | 0.6 | (1.1) | 3.8 | (6.1) | 3.6 | (4.4) | 39.8 | <0.001 |
| **FIM+FAM efficiency** | 3.4 | (4.8) | 3.8 | (4.3) | 1.1 | (1.2) | 5.4 | (6.1) | 5.7 | (6.8) | 36.6 | <0.001 |
| **NPDS/NPCNA** | | | | | | | | | | | | |
| **Admission** | | | | | | | | | | | | |
| **NPDS** | 34.9 | (19.7) | 24.7 | (17.5) | 16.5 | (15.2) | 28.9 | (16.8) | 29.4 | (15.9) | 51.2 | <0.001 |
| **Care hours/wk** | 47.3 | (9.7) | 38.7 | (21.0) | 27.8 | (22.0) | 43.2 | (20.7) | 44.2 | (18.9) | 26.6 | <0.001 |
| **Care costs** | £1,733 | (£991) | £1,205 | (£910) | £936 | (£781) | £1,578 | (£975) | £1,537 | (£869) | 49.0 | <0.001 |
| **Discharge** | | | | | | | | | | | | |
| **NPDS** | 26.0 | (21.5) | 16.8 | (17.1) | 10.2 | (12.8) | 19.3 | (16.0) | 17.2 | (16.0) | 57.1 | <0.001 |
| **Care hours/wk** | 35.2 | (24.1) | 27.5 | (21.9) | 19.6 | (17.6) | 30.5 | (20.8) | 28.3 | (21.4) | 20.8 | <0.001 |
| **Care costs** | £1,269 | (£1,081) | £851 | (£885) | £603 | (£650) | £1,068 | (£948) | £942 | (£914) | 31.1 | <0.001 |
| **Change** | | | | | | | | | | | | |
| **NPDS** | 8.9 | (12.4) | 7.9 | (11.9) | 6.3 | (12.0) | 9.6 | (13.1) | 12.1 | (12.7) | 28.9 | <0.001 |
| **Care hours/wk** | 11.6 | (16.6) | 11.2 | (14.9) | 8.3 | (18.2) | 12.7 | (16.5) | 15.9 | (16.6) | 24.4 | <0.001 |
| **Care costs** | £441 | (£808) | £356 | (£707) | £320 | (£764) | £511 | (£873) | £596 | (£798) | 20.1 | <0.001 |
| **Time to offset cost of rehabilitation** | 31.3 | (193.9) | 20.4 | (118.0) | 87.2 | (270.4) | 15.4 | (104.0) | 14.5 | (136.4) | 2.4 | 0.0.62 |

One-way ANOVAs demonstrated significant between-group differences for different levels of service for all parameters except the time to offset costs of rehabilitation due to the large spread and close 95% confidence intervals for this parameter (see *Figure 68*). Given the large size of the sample, even small differences may reach statistical significance. To adjust for this, only P values of <0.001 after Bonferroni correction were taken as significant.

**Rehabilitation Complexity:**

The length of stay closely mirrored the distribution of RCS-E scores on admission as shown in *Figure 69*.

- Rehabilitation complexity was significantly higher for Level 1 services than all other service levels. Significant differences were also seen between levels 2a and 2b, but not between Levels 1b and 2a.
- Changes scores were significantly smaller in the Level 1 than in the Level 2 services, but there were no significant differences between the Level 1a and 1b or the Level 2a and 2b services.

**FIM+FAM:**

- On admission, FIM+FAM scores were significantly lower for the Level 1a than all other service levels (p<0.001). This difference was reflected both in the motor and the cognitive subscales and in the total score. However, the differences between other service levels were not significant.
- The above findings still held true at discharge.
- Change in total FIM+FAM score was significantly higher for the Level 2a services, than for Level 1a and 1b services, but no significant differences were seen between the other service levels.
- FIM and FIM+FAM efficiency were significantly higher in the Level 2 than the Level 1 services. No significant differences were seen between 1a and 1b, or between 2a and 2b services.

**NPDS/NPCNA:**

- On admission, the total NPDS score and NPCNA-estimated care hours and costs were significantly lower in service Level 1b than in the other service levels, but no significant differences were seen between Level 1a and the Level 2 services.
- On discharge, dependency and costs remained higher in the Level 1a services with no significant differences between other service levels.
- Change in NPDS, and care hours were significantly higher in the Level 2b services than other service levels.
- The time to offset costs was not significantly different between any of the service levels.

**Discussion**

In summary, substantial differences were expected and indeed found between the Level 1c services and other service types with respect to motor and cognitive independence scores. Patients in cognitive behavioural rehabilitation settings are more likely to be independently mobile (often referred to as 'the walking wounded') and self-caring to a substantial degree. Scales such as the NPDS, emphasise physical dependency. The FIM+FAM is designed to provide greater capture of cognitive and psychosocial function than the FIM, but nevertheless, at least half of its items address physical (motor) function. Moreover, cognitive recovery is known to proceed at a slower rate than physical recovery. Therefore it was to be expected that the Level 1c services would appear 'less efficient' on measures of efficiency that rely on FIM+FAM or NPDS scores. Similarly, the Level 1b services take a mixed caseload of patients with either physical or cognitive behavioural problems, and so their mid-way position between the Level 1a and Level 1c profiles was also expected.

Some differences were seen in measures of efficiency between Level 1 and 2 services. Level 2 services were generally more 'efficient' in terms of FIM and FIM+FAM efficiency, which was to a large extent explained by the shorter lengths of stay in these services, reflecting their less complex case-load. However, when the time to offset costs was examined, there were no significant differences between the four service levels.

## *Predictors of the length of stay and cost of in-patient rehabilitation*

*Table 114* shows the bivariate correlations for possible predictive factors for length of stay and episode costs.

**Table 114: Bivariate Pearson correlations for factors related to length of stay and episode costs**

| | | | | | | FIM+FAM | | | |
|---|---|---|---|---|---|---|---|---|---|
| | **Length of stay** | **Episode cost** | **Age** | **Time since onset** | **Motor** | **Cognitive** | **Total** | **NPDS** | **Care hours/wk** |
| **Episode cost** | **0.98** | | | | | | | | |
| **Age** | -0.10 | -0.12 | | | | | | | |
| **Time since onset** | -0.08 | -0.08 | 0.01 | | | | | | |
| **Admission scores for:** | | | | | | | | | |
| **FIM+FAM** | | | | | | | | | |
| **Motor** | **0.40** | **-0.38** | -0.13 | -0.01 | | | | | |
| **Cognitive** | **0.38** | **-0.39** | 0.02 | 0.01 | **0.50** | | | | |
| **Total** | **0.45** | **-0.45** | 0.07 | 0.05 | **0.88** | **0.85** | | | |
| **NPDS** | | | | | | | | | |
| **Total** | **0.40** | **0.37** | 0.09 | 0.05 | **-0.82** | **-0.58** | **-0.81** | | |
| **Care hours** | **0.35** | **0.32** | 0.16 | 0.12 | **-0.82** | **-0.42** | **-0.73** | **0.88** | |
| **Care costs** | **0.32** | **0.30** | 0.14 | 0.02 | **-0.69** | **-0.48** | **-0.68** | **0.86** | **0.84** |

*All significant at p<0.001. Correlations of r ≥0.3 are marked in bold*

Given the size of the dataset, all correlations reached statistical significance, but factors with a correlation of $r \geq 0.30$ were entered into the stepwise multiple regression analyses. Table 115 summarises the regression models for predictors of length of stay and episode costs.

When length of stay was the dependent variable, the admission FIM+FAM motor score was entered first, accounting for 15.5% of the variance. FIM+FAM cognitive score and NPDS score were also included in the model, together accounting for a further 4%, so that the three-factor model predicted 19.7% of the variance in length of stay.

With episode cost as the dependent variable, the FIM+FAM cognitive scale was entered first, accounting for 13.9% of the variance, followed by total NPDS and FIM+FAM motor score, so that the total model accounted for 18.5% of the variance.

**Table 115: Stepwise multiple regression models for length of stay and Episode costs**

| **Dependent variable: Length of stay** | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Variables** | **Multiple R** | **Adjusted R square** | **B** | **95% CI** | **Beta** | **T** | **Significance** |
| **Admission** | | | | | | | |
| **FAM-Motor** | 0.394 | 0.155 | -0.48 | -0.57, -0.38 | -0.22 | -10.1 | <0.001 |
| **FAM-Cog** | 0.442 | 0.195 | -0.51 | -0.58, -0.43 | -0.21 | -13.8 | <0.001 |
| **NPDS-Total** | 0.445 | 0.197 | -0.28 | 0.12, 0.44 | -0.08 | 3.4 | 0.001 |
| **Dependent variable: Episode cost** | | | | | | | |
| **Variables** | **Multiple R** | **Adjusted R square** | **B** | **95% CI** | **Beta** | **T** | **Significance** |
| **Admission** | | | | | | | |
| **FAM-Cog** | 0.372 | 0.139 | -278 | -314, -242 | -0.24 | -15.3 | <0.001 |
| **FAM-Motor** | 0.430 | 0.185 | -224 | -270, -179 | -0.22 | -9.7 | <0.001 |
| **NPDS-Total** | 0.431 | 0.185 | 80 | 1.8, 157 | -0.05 | 2.0 | 0.045 |

*Excluded variables: NPCNA-estimated care hours, NPCNA estimated costs*

## *Predictors of efficiency and cost-efficiency*

Simple correlations to potential factors relating to parameters of efficiency are shown in *Table116.*

**Table 116: Bivariate Pearson correlations for factors related to length of stay and episode costs**

| | FIM efficiency | FAM efficiency | Cost efficiency | Time to offset costs |
|---|---|---|---|---|
| FIM efficiency | **1** | **0.96**** | **-0.42**** | **0.02** |
| FAM efficiency | 0.96** | 1 | -0.42** | 0.02 |
| Cost efficiency | -0.42** | -0.42** | 1 | -0.04* |
| Time to offset costs | 0.02 | 0.02 | -0.04* | 1 |
| Age on admission | -0.04** | -0.05** | 0.03 | 0.03 |
| Length of stay (days) | **-0.37**** | **-0.38**** | 0.22** | -0.08** |
| Service Level | 0.16** | 0.16** | -0.12** | 0.04* |
| **Admission scores** | | | | |
| RCS-E v 12 | -0.10** | -0.12** | 0.06** | -0.01 |
| NIS Total | -0.22** | -0.24** | 0.13** | 0.001 |
| Barthel Total | 0.06** | 0.13** | -0.03* | -0.04** |
| PCAT Total | -0.16** | -0.15** | 0.08** | 0.01 |
| FIM-Motor | 0.04** | 0.12** | -0.10** | -0.03 |
| FIM-Cog | 0.09** | 0.06** | -0.12** | 0.003 |
| FIM Total | 0.07** | 0.12** | -0.12** | -0.02 |
| FAM-Motor | 0.04** | 0.11** | -0.10** | -0.03 |
| FAM-Cog | 0.10** | 0.06** | -0.13** | 0.01 |
| FAM Total | 0.08** | 0.10** | -0.13** | -0.01 |
| NPDS Total | -0.13** | -0.16** | -0.04* | 0.04* |
| Care hrs/wk | -0.10** | -0.14** | -0.07** | 0.04** |
| Care Costs/wk | -0.12** | -0.15** | -0.21** | 0.05** |
| **Change scores** | | | | |
| NIS Total | **0.31**** | **0.30**** | -0.18** | 0.01 |
| Barthel Total | **0.50**** | **0.45**** | **-0.31**** | -0.01 |
| FIM-Motor | **0.59**** | **0.52**** | -0.23** | -0.01 |
| FIM-Cog | **0.37**** | **0.43**** | -0.12** | -0.04* |
| FIM Total | **0.61**** | **0.56**** | -0.22** | -0.02 |
| FAM-Motor | **0.61**** | **0.54**** | -0.24** | -0.02 |
| FAM-Cog | **0.40**** | **0.49**** | -0.14** | -0.05** |
| FAM Total | **0.59**** | **0.59**** | -0.22** | -0.04* |
| NPDS Total | **-0.33**** | **-0.30**** | **0.44**** | -0.003 |
| Care hrs/wk | **-0.36**** | **-0.32**** | **0.48**** | 0.002 |
| Care Costs/wk | -0.24** | -0.22** | **0.69**** | -0.02 |

*** Significant at the 0.01 level; * Significant at the 0.05 level; (2-tailed)*

Once again, given the size of the dataset, nearly all correlations reached statistical significance, but factors with a correlation of $r \geq 0.30$ were entered into the stepwise multiple regression analyses. It should be noted that factors such as service level, age, rehabilitation complexity, rehabilitation needs and the level of functional independence on admission all failed to show sufficient correlation with the efficiency parameters to support their inclusion in the regression analysis. Table 117 summarises the regression models for predictors of FIM and FIM+FAM efficiency.

When FIM efficiency was the dependent variable, the FIM+FAM motor change score was entered first, accounting for 35.1% of the variance. Change scores for FIM+FAM cognitive, NPDS care hours per week and Neurological Impairment Scale were also included in the model, but together accounted only for a further 3.5%, so that the final model accounted for model predicted 38.6% of the variance in FIM-efficiency.

When FIM+FAM efficiency was the dependent variable, the FIM+FAM motor change score was again entered first, accounting for 27.4% of the variance. Change scores for FIM+FAM cognitive, NPDS care hours per week and Neurological Impairment Scale were again included in the model, together accounting for a further 9.4%, so that the final model predicted 36.8% of the variance in FIM-efficiency.

When NPCNA cost efficiency (reduction in weekly cost of care / length of stay) was the dependent variable, change in NPCNA-estimated care hours per week was entered first, accounting for 24% of the variance. Change in NPDS score and FAM Motor score together accounted for a further 3.2%, so that the final model accounted for 27% of the variance.

Regression modelling was not completed for Time to offset the costs of rehabilitation as none of the bivariated correlations exceeded r=0.3.

**Table 117: Stepwise multiple regression models for FIM and FIM+Fam efficiency**

| **Dependent variable: FIM efficiency** | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Variables** | **Multiple R** | **Adjusted R Square** | **B** | **95%CI** | **Beta** | **T** | **Significance** |
| **Change score:** | | | | | | | |
| **FAM-Motor** | 0.593 | 0.351 | 0.09 | 0.08, 0.10 | 0.54 | 13.2 | <0.001 |
| **FAM-Cog** | 0.612 | 0.374 | 0.06 | 0.04, 0.08 | 0.24 | 7.1 | <0.001 |
| **NPDS-Total** | 0.617 | 0.379 | 0.05 | 0.02, 0.07 | 0.17 | 3.8 | <0.001 |
| **Care hours/wk** | 0.622 | 0.384 | -0.03 | -0.05, -0.01 | -0.13 | -2.8 | 0.004 |
| **NIS** | 0.624 | 0.386 | -0.05 | -0.09, -0.004 | -0.07 | -2.2 | 0.030 |
| *Excluded variables: NPCNA-estimated costs* | | | | | | | |
| **Dependent variable: FIM+FAM efficiency** | | | | | | | |
| **Variables** | **Multiple R** | **Adjusted R Square** | **B** | **95%CI** | **Beta** | **T** | **Significance** |
| **Change score:** | | | | | | | |
| **FAM-Motor** | 0.525 | 0.274 | 0.11 | 0.09, 0.13 | 0.38 | 13.2 | <0.001 |
| **FAM-Cog** | 0.592 | 0.349 | 0.16 | 0.14, 0.19 | 0.39 | 7.1 | <0.001 |
| **NPDS-Total** | 0.599 | 0.357 | 0.09 | 0.05, 0.14 | 0.20 | 3.8 | <0.001 |
| **Care hours/wk** | 0.606 | 0.364 | -0.06 | -0.09, -0.03 | -0.15 | -2.8 | 0.001 |
| **NIS** | 0.610 | 0.368 | -0.09 | -0.16, -0.02 | -0.09 | -2.2 | 0.10 |
| *Excluded variables: NPCNA-estimated costs* | | | | | | | |
| **Dependent variable: NPCNA-estimated Cost efficiency** | | | | | | | |
| **Variables** | **Multiple R** | **Adjusted R Square** | **B** | **95%CI** | **Beta** | **T** | **Significance** |
| **Change score:** | | | | | | | |
| **Care hours/wk** | 0.491 | 0.240 | 0.57 | 0.45, 0.70 | 0.38 | 9.0 | <0.001 |
| **NPDS-Total** | 0.504 | 0.252 | 0.40 | 0.25, 0.55 | 0.39 | 5.2 | <0.001 |
| **FAM-Motor** | 0.521 | 0.269 | 0.19 | 0.11, 0.14 | 0.27 | 4.6 | <0.001 |

*Excluded variables: FAM-Cognitive change, NIS change.*

**In conclusion:**

For the most part the principal predictors of efficiency and cost-efficiency could have been predicted as they were the constituent parameters that make up the calculation. However, it was interesting that change in the FIM+FAM and NPDS were stronger predictors of FIM efficiency than change in the FIM itself. Essentially none of the parameters examined gave any meaningful predictor of the time to offset the costs of rehabilitation, which emphasises that this is a considerably more complex parameter, dependent on many more influences than those tested here.

# Benchmarking on quality and outcomes

In addition to providing data for research, the UKROC database also provides information for national benchmarking of quality indicators, and the opportunity for services to compare their performance against nationally agreed standards.

Routine reporting and feedback has now been established in line with the National Outcomes Framework. The NHS England service specification for Specialist Rehabilitation for patients with Highly Complex Needs[44] sets out a set of core standards with reference to the five domains of the National Outcomes Framework. These are listed in *Table 118.*

**Table 118: Core standards from the NHSE service specification and UKROC reporting**

| | | |
|---|---|---|
| **Domain 1** | **Preventing people from dying prematurely** | |
| **Domain 2** | **Enhancing quality of life for people with long-term conditions** | |
| **A** | All patients have a defined set of person-centred goals and levels of goal attainment (GAS) recorded at discharge from the programme | UKROC provides the facility to record goals and automated calculation of GAS |
| **B** | All patients staying longer than 6 months have documented evidence that the extended length of stay is justified by the clinical gains | UKROC routinely identifies all admissions >180 days and provides the facility to record clinical gains for any extension period |
| **Domain 3** | **Helping people to recover from episodes of ill-health or following injury** | |
| **A** | Patients will be assessed within 10 days of referral | UKROC routinely reports |
| **B** | Patients will be admitted to a facility assessed as being best to meet their needs within 6 weeks of being fit for transfer | Response times for:<br>• Referral to assessment<br>• Assessment to admission |
| **C** | Tertiary specialised services will have a caseload complexity of >85% Category patients and >67% RCS-E $\geq$ 11) | Caseload complexity data<br>• Median RCS-E<br>• %RCS-E $\geq$ 11<br>• Proportion of category A patients |
| **D** | All patients will have at least one of the approved standardised measures of function recorded on admission and discharge from the programme | All approved functional outcomes including the UK FIM+FAM and FIM, NPDS and Barthel Index. |

| Domain 4 | Ensuring people have a positive experience of care | |
|---|---|---|
| A | Patients and families are generally satisfied with their care. Constructive feedback is recorded, reviewed and acted upon | Most trusts have their own patient satisfaction questionnaires in place.<br>An adapted tool based on the 'Talking Mats' system has been developed as part of the programme to support patients with cognitive / communication problems to provide feedback |
| **Domain 5** | **Treating and caring for people in safe environment and protecting them from avoidable harm** | |
| A | All services meet at least the minimum standards for safe and effective staffing levels as laid down in the BSRM standards* | UKROC collates annual reports on staffing profiles |
| B | There is no needless harm from VTE – an appropriate VTE decision tree is recorded for all patients | Most trusts have their VTE compliance systems in place. |

** Specialist neuro-rehabilitation services: providing for patients with complex rehabilitation needs. British Society of Rehabilitation Medicine. London*[17]

From April 2014, UKROC has provided regular quarterly reports to providers and commissioners with activity analysis and benchmarking (response times, outcomes and cost-efficiency) for each registered service against the core standards listed under domain 3. Performance is recorded year by year for the last 3 years, and compared with the mean performance of the other services within the same level (1a, 1b, 1c an 2a). Reports are first sent to the provider to check any anomalies and correct any potential inaccuracies, before sharing with the commissioners.

An exemplar quarterly report is given in *Appendix 8.1*.

The UKROC software also includes the facility to record Goal Attainment Scaling. It also identifies patients staying beyond 180 days and reports these back to the providers/commissioners. Local Area teams vary somewhat in their enforcement of the 180-day limit and the requirement in the specification for pre-approval of any extension, as well as post discharge reporting to confirm that the expected gains were achieved. Some areas do have robust approval and monitoring systems in place and, where this is the case, the UKROC software can be used to support forecasting of outcomes (including value for money in terms of predicted savings in the cost of ongoing care in the community).

# Comparison of performance

In addition to the routine reporting back to individual services, the UKROC team has provided regular regional road-shows and workshops for groups of local providers and their NHSE and CCG commissioners.

- In the early stages, road-shows were used to support local engagement by training clinical staff in how to apply the UKROC measures, enter the data etc.
- They were also used to engage managers and commissioners, explaining how the UKROC data could help them to plan and evaluate their service and to make the case for further development of specialist rehabilitation services.[216, 217]

- As the databases became established, they were used to provide feedback, helping local providers and commissioners to understand the differences between their various services and the reasons for them.

Specialist rehabilitation services come in a range of different types and models. It is not expected that they should all be identical. It is, however, expected that they should different in a way that is predictable from the nature of their practice and clinical caseload. Comparative data can be helpful to describe those differences that resonate with clinical experience. But they also provide the teams with an opportunity to reflect on their own practice and to consider whether all of the variation is explainable on clinical grounds, or whether there are changes that could be made to improve the efficiency of daily practice and/or the quality of care provided to patients.

## *Exemplar comparison of services*

Below we present a comparative evaluation of the London NHSE-commissioned services that was shared with NHSE and CCG commissioners across the London region in July 2014, and also with the Pan London Stroke forum in September 2014.

London has eight NHSE commissioned providers which are currently providing a range levels of services, signposted as shown in Table 119. One of the providers (Blackheath), offers two distinct units – one Level 2a and the other Level 1c.

**Table 119: Specialised rehabilitation services in London commissioned by NHSE-London**

| Service name | Area | Level | N episodes |
|---|---|---|---|
| Regional/Hyper-acute Rehabilitation Unit (RHRU), Northwick Park Hospital | NW | HA/1a | 324 |
| Royal Hospital for Neurodisability, Putney | SW | 1a | 313 |
| Regional Neurorehabilitation Unit (RNRU), Homerton | NE | 1b | 216 |
| National Rehabilitation Unit, University College London Hospitals | NC | 2a/b | 256 |
| Frank Cooksey Rehabilitation Unit, King's College Hospital | SE | 2a/b | 204 |
| Wolfson Rehabilitation Unit, St George's Hospital | SW | 2a/b | 138 |
| Thames Brain Injury Unit, Blackheath | SE | 1c | 96 |
| High Dependency Neurorehabilitation Unit, Blackheath | SE | 2a | 99 |
| Lishman Unit, South London and Maudsley | SE | 1c | 37 |

Performance data in relation to the standards were extracted for all episodes discharged between April 2010 and 31st March 2014. The number of episodes for each unit is given in *Table 113 .*

In the comparative reports below, the individual units are identified only as single digits for privacy. They are included in the graphs in a consistent order, although this is not the same order as the table above.

## *Response times:*

*Figures 72 and 73* show the mean waiting between a) referral and assessment by the receiving unit (standard ≤10 days) and b) between assessment and admission (standard ≤42 days). The mean response times are summarised in *Table 120.*

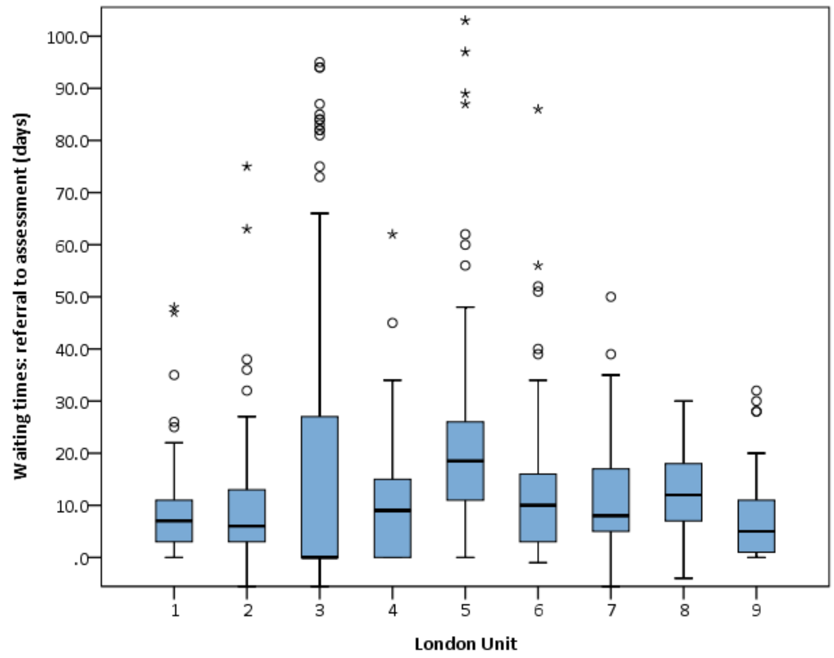



**Figure 71:** **Response times: a) Referral to assessment (Standard ≤10 days)**

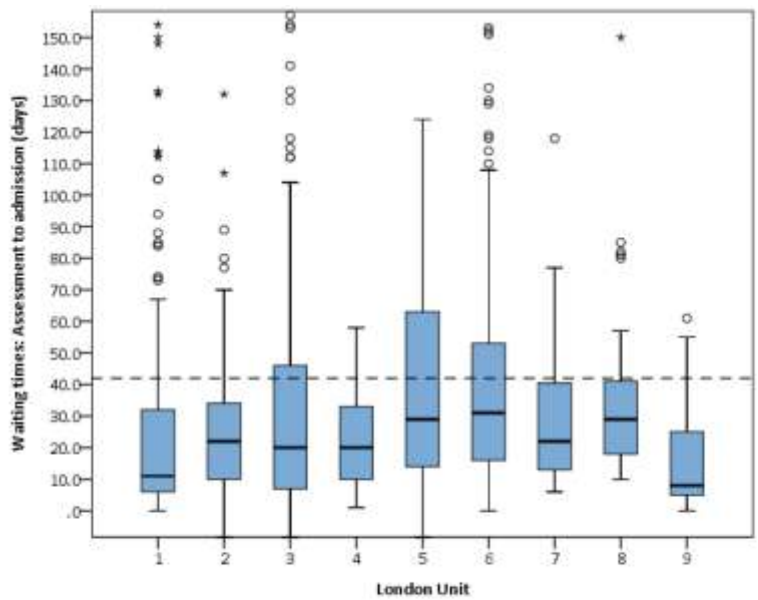



**Figure 72:** **Response times: b) Assessment to admission (Standard ≤42 days)**

**Table 120: Comparative response times – mean (SD) days**

| Standard | Referral to assessment ≤10 days | | Assessment to admission ≤42 days | | Total referral to admission ≤52 days | | Length of stay ≤180 days | | Delayed discharge ≤14 days | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Unit** | **Mean** | **(SD)** | **Mean** | **(SD)** | **Mean** | **(SD)** | **Mean** | **(SD)** | **Mean** | **(SD)** |
| 1 | 8.3 | (10.3) | 23.0 | (29.1) | 31.3 | (31.2) | 102.8 | (52.5) | 1.2 | (4.2) |
| 2 | 11.1 | (32.0) | 22.1 | (36.9) | 33.2 | (22.7) | 73.2 | (47.3) | 3.7 | (21.2) |
| 3 | 22.5 | (53.3) | 42.0 | (68.5) | 64.5 | (91.5) | 43.5 | (27.5) | -0.3 | (8.9) |
| 4 | 10.2 | (9.9) | 22.1 | (13.5) | 31.9 | (18.2) | 79.7 | (31.4) | -0.5 | (6.3) |
| 5 | 17.2 | (37.4) | 41.2 | (48.6) | 55.5 | (63.0) | 131.6 | (57.9) | 2.4 | (28.0) |
| 6 | 11.3 | (10.4) | 38.9 | (33.2) | 50.3 | (34.6) | 160.6 | (75.0) | 14.4 | (27.1) |
| 7 | 18.8 | (499.6 | 23.6 | (52.0) | 42.6 | (29.4) | 136.2 | (82.0) | 0.6 | (8.7) |
| 8 | 12.6 | (7.4) | 38.0 | (30.8) | 51.3 | (41.0) | 122.5 | (63.4) | 4.2 | (21.1) |
| 9 | 8.2 | (9.3) | 20.3 | (31.2) | 28.4 | (34.2) | 157.7 | (81.4) | 17.5 | (56.7) |
| **All** | **13.3** | **(30.2)** | **31.9** | **(43.4)** | **45.0** | **(51.9)** | **107.0** | **(69.1)** | **4.1** | **(20.5)** |

Mean figures not meeting the standard are shown in red

Of the nine services, five had waiting times for assessment above the standard. However, mean waiting times from assessment to admission all fell within the standard, so that the overall average waiting times from referral to admission were within the standard (≤52 days) for all but one service.

Nevertheless, as *Figure 73* illustrates, there were a significant number of substantial outlying cases, some of whom waited 6 months or more between initial referral and assessment.

- In some cases this was because the patient continued to have unstable medical needs requiring continued stay in the acute setting.
- In other cases it was simply due to lack of bed availability in the receiving unit, which meant that early assessment was of low priority, as the patient could not be admitted anyway.

The variation provided an opportunity to review differences in practice. Units 1 and 9, which had amongst the lowest waiting times for assessment, both have consultants who regularly perform outreach visits to assess patients in the acute services. Unit 5 only offered out-patient assessments at the unit itself, which (a) provided very limited capacity and (b) required the patient to be fit enough to travel to the unit. After reviewing this practice, Unit 5 enhanced its service by providing outreach assessments as well, to reduce the waiting time for assessment.

## *Length of stay*

The service specification covers in-patient rehabilitation episodes of up to 180 days for patients with category A needs. Figure 74 shows the distribution of Length of stay (LOS) for the nine services. It is not necessarily expected that all units would have the same LOS profile. However, we did expect that variations would be explainable by the type of service and the nature of the caseload.

Units 1-4 all had relative short mean length of stay (ranging from 43 to 103 days (see Table 121). The remaining services all had substantially longer lengths of stay. Two of these (units 8 and 9) were cognitive behavioural rehabilitation services where LOS is expected to be longer and units 6 and 7 were known to provide relative slow-stream services for patients with complex physical disability (see data on FIM+FAM scores below). Thus the variations were largely as expected with the exception of unit 5.

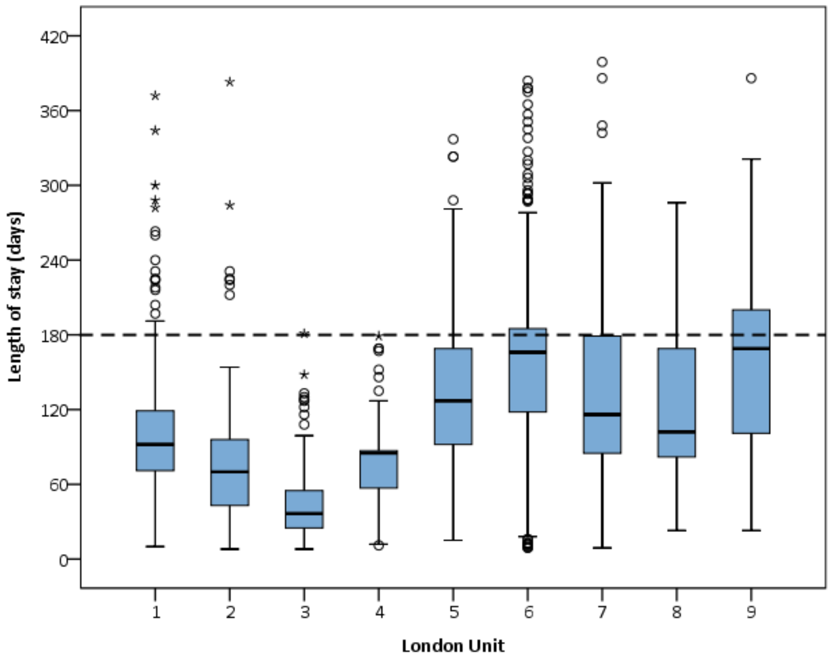



**Figure 73:** **Length of stay – compared between units (Standard ≤180 days)**

## *Functional outcome*

*Figure 74* shows the total FIM+FAM scores on admission (in blue) and discharge (in green) across the nine units. Once again variation was expected. For example, Unit 6 takes a significant proportion of patients in vegetative and minimally conscious states, many of who will remain profoundly disabled – hence the low FIM+FAM scores for this service are not unexpected.

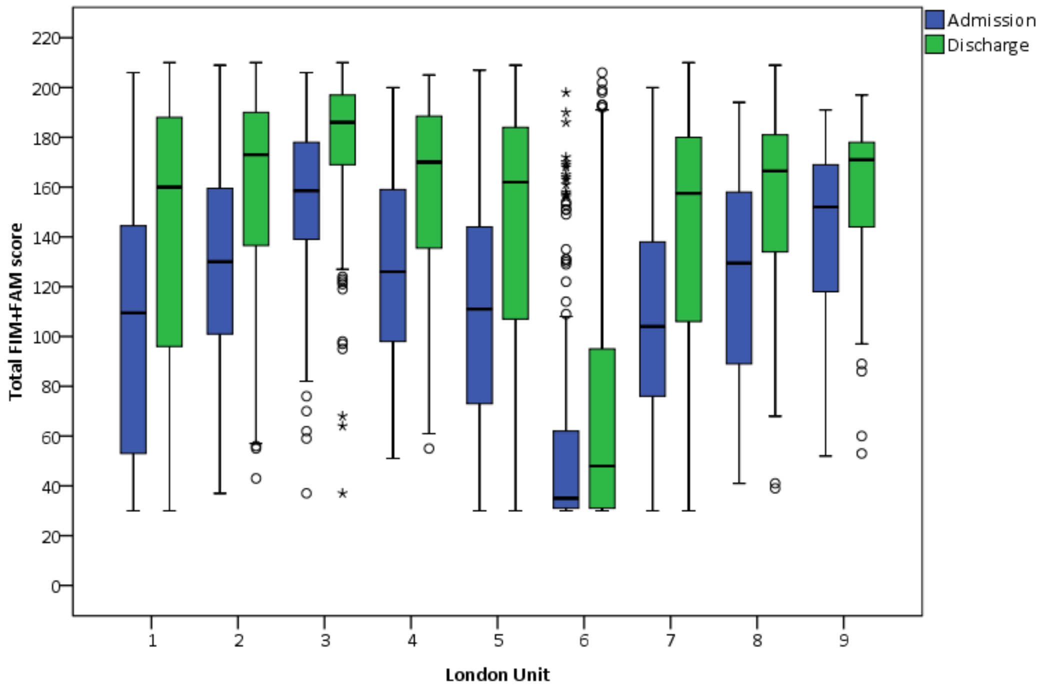



**Figure 74:** **Functional independence: Total FIM+FAM scores from admission to discharge**

The total FIM+FAM scores give a general picture of the heterogeneity of the case loads across the nine units, but Figure 67 shows the composite FAM-splat for the different services and provides an even clearer picture of the different profiles, which appear to group as follows:

- Units 1, 5, 6 and 7 (Figure 75) generally take a complex physical caseload – as mentioned above, while Unit 6 has a particular emphasis on patients with profound brain injury.
- Units 8 and 9 (Figure 76) manage patients with cognitive behavioural problems with relatively less physical disability.
- Units 2, 3 and 4 (Figure 77) take a mixed caseload. Unit 3 has a particular interest in patients with multiple sclerosis and functional deficits, admitting a higher proportion of patients from the community who have relatively intact communication skills.


**Figure 75: Composite FAM-splats for units 1, 5 6 and 7 – complex physical disability**

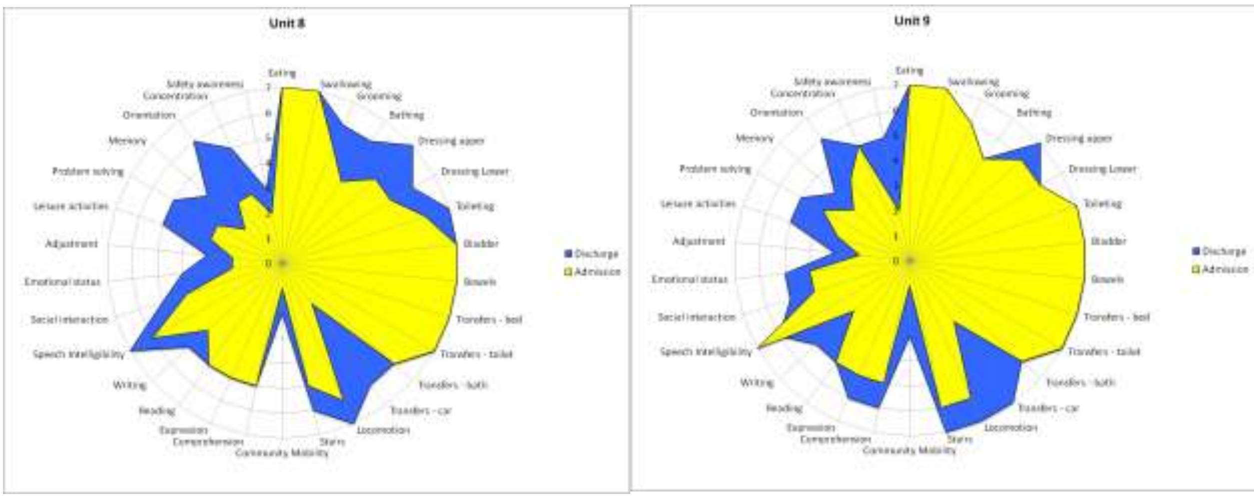



**Figure 76:** **Composite FA-Splats for units 8 and 9 – cognitive behavioural disability**

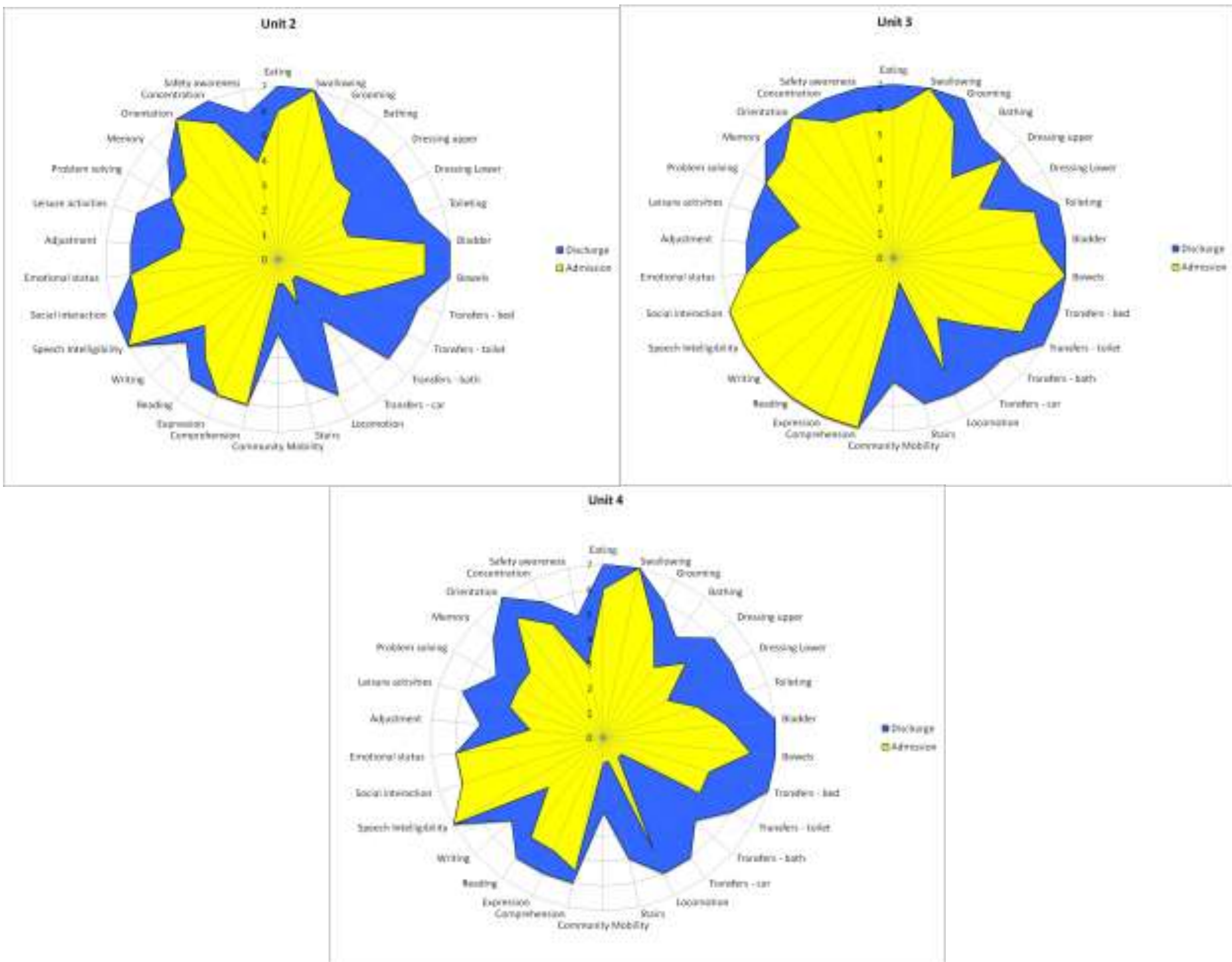



**Figure 77:** **Composite FAM-Splats for units 2, 3 and 4 – mixed caseload**

## *Cost efficiency:*

*Table 121* shows the mean values for length of stay, episode costs, change in functional independence and care costs in the community, with efficiency indices calculated on a unit basis. Episode costs were calculated from the actual reported cost per occupied bed day, as reported to UKROC by each of these nine units (rather than the mean service level costs used in *Table 117* above). At service level:

- Mean episode costs ranged from £18,044 – 67,012.
- The mean reduction in on-going care costs per week as a result of these admissions was £90-£522.
- The mean time to offset the cost of rehabilitation ranged from 16.6 to 157.3 months.
- The patients with the most profound physical and cognitive problems had longer lengths of stay (mean 131-157) days which impacted on the indices of cost-efficiency.

**Table 121: Comparative analysis of the London Units for mean values for change in functional independence and care costs between admission and discharge, and indices of efficiency**

| Unit | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| Length of stay | 102 | 73 | 43 | 79 | 131 | 160 | 136 | 122 | 157 |
| Episode cost* | £57,979 | £42,887 | £18,044 | £25,268 | £53,284 | £53,492 | £36,358 | £46,040 | £67,012 |
| **FIM+FAM** | | | | | | | | | |
| Admission | 103.7 | 127.2 | 153.5 | 127.2 | 106.3 | 54.2 | 100.2 | 124.7 | 142.2 |
| Change | 36.7 | 33.3 | 24.4 | 32.5 | 33.8 | 17.5 | 32.9 | 29.3 | 17.7 |
| **NPDS** | | | | | | | | | |
| Admission | 28.8 | 19.9 | 10.7 | 21.1 | 28.7 | 49.2 | 25.8 | 20.9 | 11.2 |
| Change | -10.4 | -5.0 | -5.7 | -7.5 | -9.2 | -3.1 | -7.9 | -9.0 | -3.7 |
| **Care costs** | | | | | | | | | |
| Admission | £1,472 | £1,075 | £562 | £1,348 | £1,412 | £2,344 | £1,393 | £1,203 | £627 |
| Change | £522 | £126 | £271 | £378 | £437 | £90 | £477 | £441 | £106 |
| **Efficiency indices (calculated on a unit basis)** | | | | | | | | | |
| **FIM+FAM efficiency | 2.5 | 3.2 | 3.9 | 2.9 | 1.8 | 0.8 | 1.7 | 1.7 | 0.8 |
| ***Cost efficiency | 35.8 | 12.1 | 44.1 | 33.5 | 23.4 | 3.9 | 24.6 | 25.3 | 4.7 |
| ****Time to offset costs (months) | 27.8 | 85.2 | 16.6 | 16.7 | 30.5 | 149.3 | 19.1 | 26.1 | 157.5 |

** Episode costs calculated from the actual cost per occupied bed day as reported to UKROC by each of the nine units*
*** FIM+FAM efficiency = Mean reduction in total FIM+FAM score / Mean length of stay (weeks)*
**** Mean cost efficiency = Mean reduction in care costs / Mean LOS (weeks)*
***** Mean time to offset costs = mean time for reduction in weekly care costs in the community to offset the cost of rehabilitation*

*Figure 78* shows the comparative FIM+FAM efficiency for the nine units, which to some extent provides a mirror image of the length of stay (see also *Figure78*)

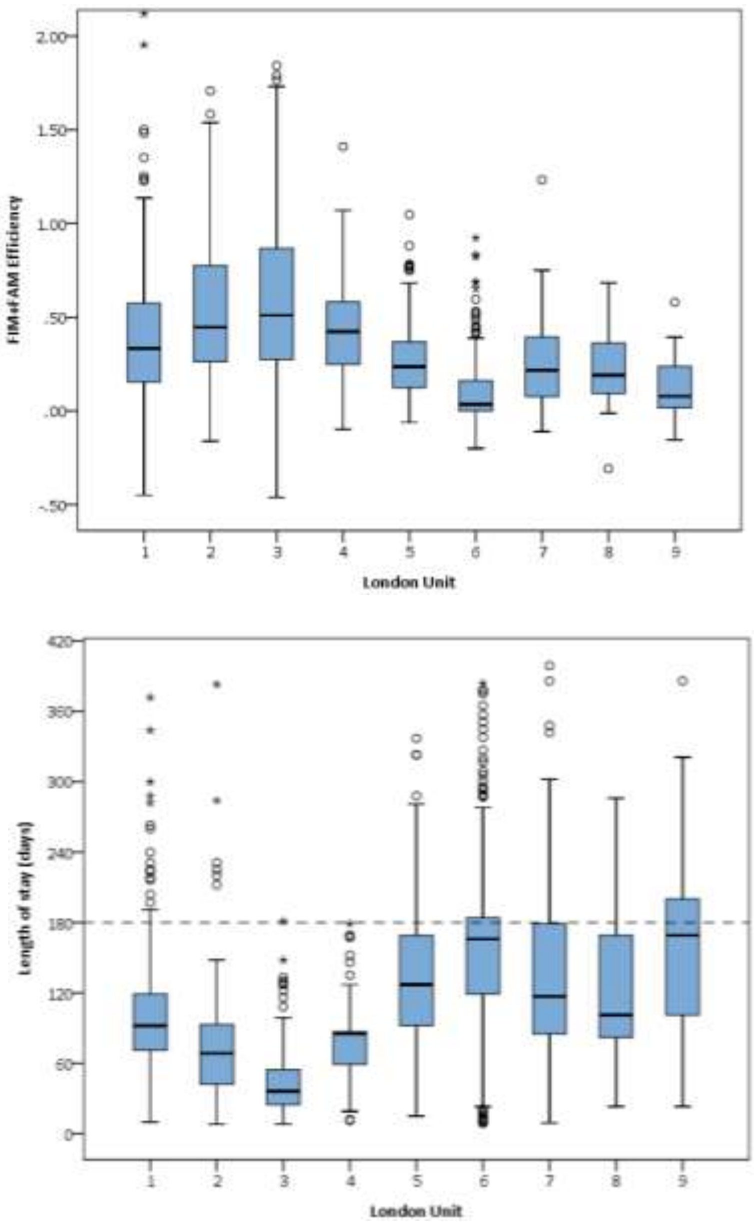



**Figure 78: Comparative FIM+FAM efficiency for the nine units compared with length of stay**

## Discussion

In conclusion, this section illustrates how the UKROC dataset has provided comparative information on the performance of other services within the region. It is expected that any region would provide a range of different levels and types of services, each with their own individual and complexity profile. Therefore it is not expected that these units should perform alike. However, it is expected that the data should provide an interpretable picture of the contribution of each unit which resonates with clinical experience.

To a large extent this was true. Nevertheless, certain anomalies were identified and raised in discussion with the provider units and, where necessary, the commissioners. Where appropriate, this resulted in recommendations for change that were acted upon - either by the provider, or by the provider in conjunction with the commissioner where additional resources were required.

These comparisons have been undertaken for regions across the country and shared with providers and commissioners – either on the basis of individual discussions or in road shows.  Space within this report does not permit the exposition of every comparison.  However, the feedback received has been very positive – the data have not only provided evidence to underpin clinical impressions, but also independent information about the resource implications for correcting identified anomalies, which is reported to have been helpful by both the local provider and commissioner teams.

# Outcomes and cost efficiency for the slow stream rehabilitation services

As noted in chapter 7, data collection from the slow stream rehabilitation services is less complete than for the Level 1 and 2 services. Nevertheless, a number of units had kindly provided data, and the analysis below summarises the information that we do have.

Within the dataset collected in the 4 years between April 2010 and March 2014, a total of 10 slow stream rehabilitation services (each with a mean length of stay >100 days) reported at least 10 cases. The resulting dataset comprised 341 episodes.

The services included a mixture of neuro-behavioural units and those catering for predominantly physical disability. As the numbers were relatively small, however, we did not attempt to subdivide the dataset by service type, but analysed all the episodes together.

The demographics were was as follows

- Males:female ratio 2.4:1
- Mean age 43 years (SD 17.2).
- Diagnoses:
    - Acquired brain injury n=235 (68.9%),
    - Spinal cord injury n=15 (4.4%),
    - Progressive conditions n=19 (5.6%),
    - Other conditions n=9 (2.6%),
    - Missing data n=63 (18.5%).
- Mean time since onset 250 (SD538) days.

The mean length of stay for the whole group was 273 days (SD 424). A total of 243 (71%) patients were discharged during the collection period: of which 130 (38% of the whole) went home and 100 (29%) went on to other institutional settings (e.g. nursing home or residential care). Eight (2%) were transferred to an acute hospital.  Figure 79 shows the distribution of length of stay within the 10 services.

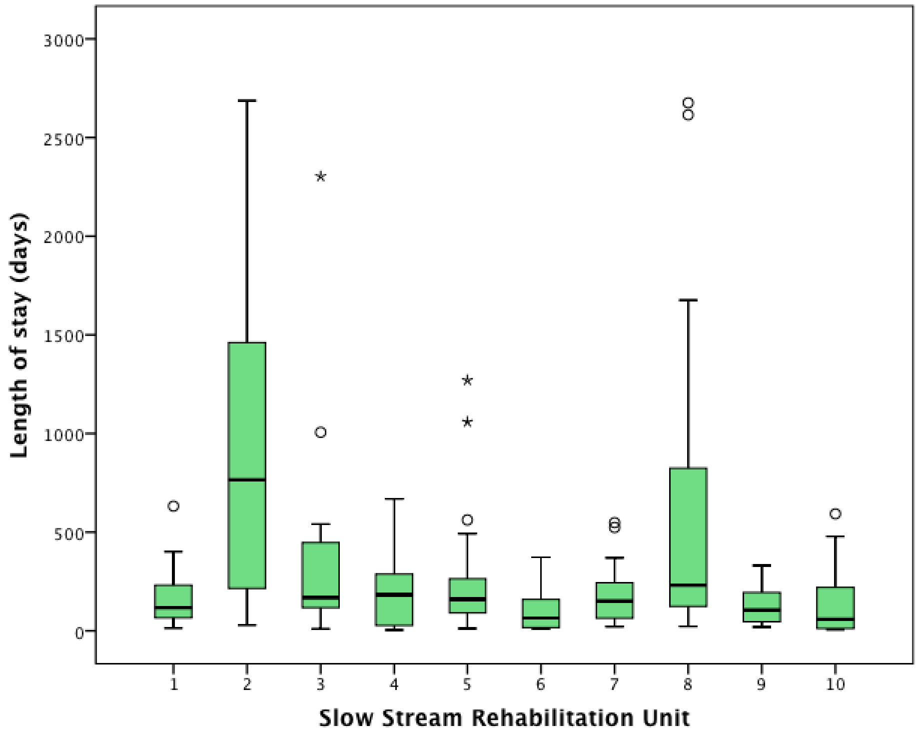



**Figure 79: Box plots of the length of stay for the 10 slow stream rehabilitation services**

It can be seen that the length of stay was quite variable between services, with two units having markedly longer lengths of stay than the others. However, these two services contributed a total of only 8 patients to the FIM+FAM and NPDS datasets summarized below, and were thus considered not to skew the data excessively.

Reporting of UKROC data was voluntary for this group, but are summarised in Table 122. Significant improvements were seen between admission and discharge on all parameters, and in fact the figures for this group are not so very different from those seen in the Level 1 and 2 services (see Table 105).

For the 148 who had a valid FIM+FAM score on admission and discharge, the mean gain in total FIM+FAM was 30.8, compared with 33.1 for the Level1 and 2 services. The mean length of stay for this subset 190 (SD 177) days, and the individually-calculated mean FIM+FAM efficiency was 1.8 (SD 2.4) compared with 4.8 for the level 1 and 2 services. This somewhat lower efficiency primarily reflects the length of stay (which was approximately 2.5 times longer for these services.

Figure 81 shows the composite FAM-splat of the median item level scores on admission and discharge.

For the 81 patients who had an NPDS score calculated on admission and discharge, the mean length of stay was 156 (SD019) days. The mean NPCNA-estimated reduction in care costs was £294 (SD 442) per week. Based on an estimated average cost of £350 per day, the mean cost of these episodes was £54,600. With a mean cost saving of £294 per week, the time taken to offset the cost of rehabilitation was 46 months (or just under 4 years).

**Table 122: Dependency and functional outcome scores on admission and discharge**

| | Admission Mean (SD) | Discharge Mean (SD) | Mean difference | 95% CIs | | t | P value* |
|---|---|---|---|---|---|---|---|
| | | | | Lower | Upper | | |
| **Rehabilitation Complexity Scale (RCS-E v 12) (n=115)** | | | | | | | |
| RCS-E | 11.1 (2.7) | 8.5 (3.1) | -2.5 | -3.1 | -1.9 | -8.3 | <0.001 |
| **Barthel Index (n=205)** | | | | | | | |
| **Total Barthel score** | 10.2 (7.3) | 13.6 (6.9) | 3.4 | 2.8 | 4.1 | 10.5 | <0.001 |
| **UK Functional Assessment Measure - UK FIM+FAM) (n=148)** | | | | | | | |
| Self-care | 31.0 (15.1) | 37.4 (14.0) | 6.5 | 5.0 | 8.0 | 8.5 | <0.001 |
| Sphincter | 9.4 (5.2) | 11.0 (4.5) | 1.6 | 1.0 | 2.2 | 5.8 | <0.001 |
| Transfers | 16.5 (10.1) | 21.5 (8.5) | 5.0 | 3.8 | 6.2 | 8.4 | <0.001 |
| Locomotion | 10.0 (6.5) | 14.1 (6.3) | 4.1 | 3.2 | 5.0 | 9.5 | <0.001 |
| Communication | 23.0 (10.0) | 26.2 (9.4) | 3.2 | 2.4 | 4.0 | 8.1 | <0.001 |
| Psychosocial | 14.8 (7.2) | 19.3 (67.5) | 4.5 | 3.6 | 5.4 | 10.2 | <0.001 |
| Cognition | 18.7 (9.6) | 24.7 (9.5) | 6.0 | 4.9 | 7.1 | 11.0 | <0.001 |
| Total Motor | 66.8 (34.7) | 83.9 (31.9) | 17.1 | 13.5 | 20.8 | 9.3 | <0.001 |
| Total Cognitive | 56.5 (24.7) | 70.1 (25.1) | 13.7 | 11.2 | 16.1 | 11.1 | <0.001 |
| **Total FIM+FAM** | 123.3 (53.1) | 154.1 (53.1) | 30.8 | 25.3 | 36.2 | 11.2 | <0.001 |
| **Functional Independence Measure (FIM) (n=167)** | | | | | | | |
| FIM-Motor | 54.0(29.3) | 68.1 (26.6) | 14.1 | 11.2 | 17.0 | 9.7 | <0.001 |
| FIM-Cognitive | 20.7 (9.2) | 25.3 (8.9) | 4.5 | 3.7 | 5.4 | 11.2 | <0.001 |
| **Total FIM** | 74.6 (34.7) | 93.4 (33.0) | 18.7 | 15.3 | 22.0 | 11.0 | <0.001 |
| **Dependency (Northwick Park Dependency Score and Care Needs Assessment – NPDS/NPCNA) (n=81)** | | | | | | | |
| Total NPDS | 25.4 (19.4) | 19.4 (19.0) | -5.9 | -8.3 | -3.5 | -4.9 | <0.001 |
| Care hrs/wk | 36.5 (20.5) | 30.4 (22.2) | -6.1 | -9.0 | -3.2 | -4.2 | <0.001 |
| Care costs/week | £1400 (943) | £1106 (940) | -£294 | -£442 | -£146 | -3.9 | <0.001 |

**two-tailed significance*

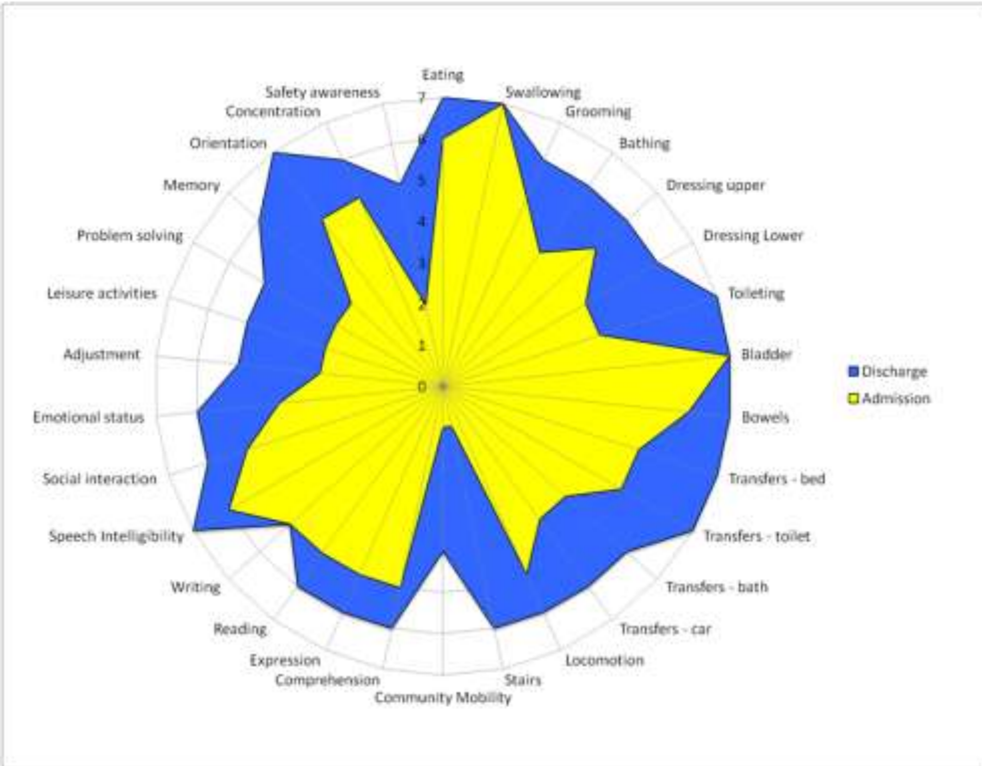



**Figure 80: Composite FAM-splat of the Median item level scores on admission and discharge**

In summary, the dataset from the slow-stream services is less complete than that from the Level 1/2 services. The information that we have, however, confirms with the extra time for rehabilitation in a slow stream setting, at least a third of these patients are able to make the transition back into the home setting. Progress is slower for these patients and the cost savings are smaller. Nevertheless, the costs of the slow-stream programme are offset by savings in ongoing care with under 4 years. Given an average age of just 43 years, this still leaves many years of life over which to recoup the benefits.

## *In summary for this Chapter*

In this part of the programme we established a nationwide database for centralised collation and analysis of case-episode data on needs, inputs, costs and person-centred outcomes from specialist neurorehabilitation services in the UK. We analysed 4 years' prospective data collection to evaluate clinical benefits and cost-effectiveness, across a range of conditions (in terms of both diagnosis and dependency) and service types. We also demonstrated how the data can be used for quality benchmarking for process and outcomes.

## *Our key deliverables were:*

1. All level 1 and 2a services to be registered and routinely submitting high quality case-episode data to the UKROC dataset.
2. Routine reporting and feedback systems established to provide contributing centres with activity analysis and benchmarking of outcomes.
3. Final analysis of cost predictors and outcome in relation to complexity and other factors (e.g. dependency) to have been undertaken on the cleaned 2013/14 dataset.

These have all been achieved.

## *Changes from the original application.*

We hoped to be able to provide a comprehensive evaluation of outcomes for slow-stream rehabilitation services in the community in the way that we have for the Level 1 and 2 services.

As noted for Chapter 7, these services were not included in the NHS England service specification, so the UKROC data collection is not mandated. Therefore although the data have been assembled and shared, this they are not so complete or robust as those from the level 1 and 2 services.

The data that we do have, however, confirm that there is a group of patients who are still able to make gains further down the line from injury, and have the potential to return home in the longer term if provided with a less intense but more prolonged slow-stream rehabilitation programme.

This area of rehabilitation is increasingly recognised as an important area for development. NHS England has become aware of the substantial on-going costs that this group of patients represent within the NHS continuing care services and the BSRM has introduced guidelines for different levels of nursing care. Our data are encouraging thus far, but further exploration is required as the dataset enlarges to identify the factors that predict which patients have potential to return home will be the subject of a follow-on grant application.

# Chapter 9: Guidance and policy development

## Objective:

Dissemination to support implementation through guidance and policy development.

## Deliverables

1. The UKROC database will be fully established as the commissioning dataset for Level 1 and 2 specialist rehabilitation services across England, and as the vehicle for administration of the multi-level weighted bed-day tariffs.
2. Data registry status for the UKROC database will have been applied for.
3. Linkages between UKROC and TARN as the basis of the National Clinical Audit (NCA) for Trauma and Complex neurorehabilitation will have been defined and the further work needed to provide the National Clinical Audit will have been scoped and applied for through the Health Quality Improvement Programme (HQIP)

Deliverables 1 and 3 have been achieved, and 2 is due for submission by September 2015. It was put back by the delayed start of the HQIP-funded programme.

This final chapter describes the dissemination and implementation of our findings through development of evidence-based policy and guidance.

## Background

This programme has evolved though a time of great change in the NHS. The Health and Social Care Act 2012 introduced the most radical re-organisation and restructuring of the commissioning / funding in the entire history of the NHS. This has provided both opportunities and challenges, as described in Chapter 2.

Throughout the programme we have shared our results through peer-reviewed publications, lectures and conference presentations to reach a wide audience - including health services managers who commission or provide rehabilitation and continuing care services, as well as clinicians (doctors, nurses and allied health professionals) in rehabilitation and related fields of practice.

Findings from our work on tool development and service episode costs have been shared with NHS England, Monitor and the Health and Social Care Information Centre to inform central policy in casemix and tariff development. Direct interaction with commissioning networks has supported the development of tools and information in formats that are useful to providers and purchasers, enabling them to use the information for clinical decision-making and to improve the services that are offered to patients.

Where we believe our approach may be relevant for wider application, we have also worked with other equivalent groups in other areas of healthcare, to share our techniques and lessons learned. These interactions have included Expert Working Panels involved in casemix and tariff development for Palliative care and Complex neurological disability in Children.

## *Utilisation of Outputs - Impact and Uptake in National policy*

The findings and developments produced in the course of this programme have been integrated into the commissioning strategy for specialised rehabilitation services as this has progressed over the seven-year life-time of this programme. They have had major impact on national policy in this area.

Since 2012, the UKROC database has provided the commissioning dataset for NHS England. From the end of this programme, we are pleased to report that the UKROC database has now been included in the NHS England commissioning portfolio for 2015/16 and going forward.

This contracting arrangement confirms the value that NHS England places on the outputs of this programme grant for the purposes of commissioning and national benchmarking.

## *NHSE Service specification for specialist rehabilitation services*

The NHS England's service specification for Specialist Rehabilitation for Patients with Highly Complex needs[44] requires that only activity reported through UKROC is eligible for commissioning under this specification. All services must be registered with UKROC and reporting the full UKROC dataset for all episodes to be eligible.

Data provided by this programme on service configuration has been used by the British Society of Rehabilitation Medicine (BSRM) to drawn up its standards for neurorehabilitation services.[17] These standards have in turned been taken up and used in the NHS-England Service specification for the designation of services.

Although service designation is ultimately a commissioning decision, UKROC provided the signposting information to enable NHS England commissioners to designate services at the appropriate level (1a, 1b, 1c, 2a, and 2b) based on caseload complexity, staffing levels and resources/expertise available within each of the service types.

## *Commissioning currencies*

The multi-level weighted bed-day currency (described in Chapter 7 of this report) was mandated as the commissioning currency for all Level 1 and 2 specialist rehabilitation services in England from April 2012.

When NHS England took over commissioning of the specialist rehabilitation services in April 2012, it inherited a number of problems:

- Although specialised commissioning groups (SCGs) had been tasked with commissioning services under the 36 original Specialised Services Definitions[14], uptake was variable and collaborative commissioning was only established in three regions around the country.
- The Department of Health's broad definition of a 'specialised' service as one that 'covered a catchment area of >1 million' tended to favour development of specialised commissioning in the more densely populated areas (e.g. London and Manchester etc), at the expense of more sparsely populated areas. To ensure equitable service provision, NHS England adopted some Level 2a (supra-district) services within its commissioning portfolio in areas where no Level 1 services existed.
- Even within the more populated areas, many services were under-commissioned. In the absence of a reliable mechanism for identifying the true costs of service provision, many so-called 'Level 1' services were commissioned at rates that would barely cover the costs of a Level 2 service.

- Unable to afford the staffing levels to take on an undiluted caseload of patients with highly complex (Category A) needs, many 'specialist' services were forced to balance their casemix by taking on a proportion of simpler (Category B) cases.
- In the first two years, a system of 'steady state commissioning' was introduced to avoid destabilising services, and NHS England commissioned both category A and B activity within the Level 2a services.

From April 2015, this 'steady state' commissioning expired, and this triggered a number of changes to ensure that NHS England commissions only the patients with highly complex needs.

- The local Clinical Commissioning Groups (CCGs) were given notice that, from April 2015, category B patients would be devolved to local commissioning. In order to assist with this transition, UKROC provided information to each NHSE Local Area Team on the estimated proportion of category B patients (PCAT<30) admitted to each of the NHS England-funded services.
- That effectively left the Level 2a services with a smaller number of centrally commissioned patients, but all of those having Category A needs.

The weighted bed day currency facilitated this transition, by enabling NHS England commissioners to purchase a smaller number of more complex patients within the same bottom line budget.

## *Tariff development*

In 2012, alongside the weighted bed day currency, the Department of Health's PbR team published a set of indicative tariffs. These were based on the costing figures provided by UKROC. NHS England recommended that commissioners should use the published indicative tariffs if at all possible, or at least to start working towards them.

The UKROC team continues to update the costing information from providers on an annual basis, and is now providing this to Monitor and NHS England, in order to keep the tariffs up to date.

During the transition phase, Monitor and NHS England were understandably preoccupied with the larger volume higher cost services, and the tariff figures lagged behind price changes as the staffing levels and complexity steadily increased (see Chapter 7). However, contact has now been re-established with the Monitor and NHS England pricing teams, and work is underway to introduce mandated tariffs for 2016/17 following price adjustment. We developed a document for commissioners and providers giving the rationale and explaining the weighted bed day currency, see Appendix 9.1. It was published by NHSE/Monitor in their 2014/15 pricing guidance.

## *Benchmarking*

Until the development of a common national dataset, there was little cross-communication or understanding of the differences between the various provider services within a given area. One major advantage of a common dataset is that it provides a consistent basis for comparison. This in turn has provided commissioners and the DoH with a better understanding of the services they purchase. Their previous view that all specialist rehabilitation services should look alike and meet the same standards has been replaced with a clearer understanding - as expressed by one commissioner:

***"I thought we were commissioning apples and pears – but I now realise that we are commissioning a whole fruit basket".***

In addition to providing comparative information in relation to quality standards, the quarterly benchmarking reports, as described in Chapter 8, have helped to identify gaps in services and support a more holistic approach to commissioning.

## *The Specialist Rehabilitation Prescription*

Within the acute care and major trauma networks, the requirement for early intensive rehabilitation is increasingly recognised as a cost-effective strategy.

The Clinical Reference Group (CRG) for Major trauma introduced the Rehabilitation Prescription for all seriously injured patients (Injury Severity Score >8), alongside the requirement for review by a consultant in Rehabilitation Medicine within 96 hours.

The British Society for Rehabilitation (BSRM) built on this recommendation by introducing a Specialist Rehabilitation Prescription (SpRP). [122, 218] The SpRP relies on UKROC tools such as the PCAT Tool, RCS-E and NPDS, in order to:

- identify those patients with complex rehabilitation needs
- define their requirements for rehabilitation, and
- direct them down the appropriate specialist Level 1/2 rehabilitation pathway in accordance with their individual needs.

The linkage between the UKROC and the Trauma and Audit Research Network (TARN) Dataset will ultimately enable data linkage, in order to track patients down the pathway from acute care into specialist rehabilitation services and thence to the community.

## *Further research programmes*

Through collaboration with research colleagues, our work as led to five further successful collaborative grant applications in the UK and overseas, as follows:

1. NIHR project grant – Health Services and Delivery Research Programme (SDO-0000183): *Evaluation of community rehabilitation service delivery in long term neurological conditions.* Principal applicants: Turner-Stokes L and Siegert RJ (£420,309, 2009-2013).
2. NIHR Programme Grant*: (*RP-PG-1210-12015). *C-CHANGE: Delivering high quality and cost-effective palliative care for patients in the last year of life and their families.* Principal applicant –Murtagh F.(£2,000,000, *2012- 2017).*
3. NIHR Programme Development Grant (RP-DG-0613-10002)*: How should the NHS deliver rehabilitation services for children after acquired brain injury?* Principal applicant – Forsyth R. (£99K, 2014-2015*).*
4. Transport Accident Commission and the Institute for Safety, Compensation and Recovery Research (ISCRR), Melbourne, Australia. *Rehabilitation after catastrophic acquired brain injury: Evaluation of process and outcomes of a specialist ABI unit and its impact on long term quality of life and community participation*. Principal applicant – Lannin N. 2014-2019.
5. Health Quality Improvement Partnership (HQIP): *National Clinical Audit for Specialist rehabilitation following major trauma.* Principal applicant: Turner-Stokes L (£900K. 2015-18).

## *In respect of 5)*

a topic proposal was successfully submitted to HQIP for a National Clinical Audit of Specialist Rehabilitation following Trauma, and was approved in 2011. After some considerable delay, it was scoped in May 2013 and invitations to tender advertised in July 2014.

A collaborative tender was successfully won, led by North West London Hospitals Trust in partnership with UKROC, the Trauma Audit and research Network (TARN) and King's College London. The project started on July 1st 2015. It will develop data linkage between the TARN and UKROC datasets to enable patients to be tracked from the major trauma centres into the Level 1/2 rehabilitation programmes, and to follow their progress.

One of the early tasks is to submit an application to the Health Research Authority for Section 251 approval to collect the identifiable data (NHS number) necessary to form the data linkage between TARN and UKROC. This application is now in development and will be submitted by September 2015.

## *Intellectual Property Policy*

The copyright for the tools is retained by Professor Turner-Stokes on behalf of Northwick Park Hospital and its governing NHS Trusts ('North West London Hospitals NHS Trust' and subsequently 'London North West Healthcare NHS Trust'). However, it is a founding principle of the UKROC programme that we make all of the tools freely available to improve the quality of patient care.

In the developmental stages of this programme grant, the latest versions of the tools were retained within the UKROC office and potential users asked to contact the UKROC team for copies. This enabled us to keep a register of users, to whom updated versions could be circulated as these were developed.

Since the UKROC database went live in 2010, all UKROC-registered units receive the updated UKROC manual, current versions of the tools and software each April. This ensures that all units submitting data are using the most recent versions. Use of the tools is monitored through the routine data reporting for each admitted episode of specialist in-patient rehabilitation. The data derived from them is collated, analysed and reported as part of our outputs for the programme.

In addition, now that the tools have largely stabilised, many of them have been placed on the UKROC website where they are directly available for download, along with training materials etc. This helps to encourage their wider use, both nationally and internationally.

## *Guidelines and standards*

Building on the research and development undertaken within this programme, a number of national standards and guidelines were developed to promote implementation of the UKROC database.

In 2013, the British Society of Rehabilitation Medicine published Core Standards for Specialist Rehabilitation following Major Trauma[218] in collaboration with the NHS England CRG for Major Trauma. These standards promote use of the Specialist Rehabilitation Prescription as described above.

In 2014, a further set of standards was published collaboration with the CRG for Critical Care. These guidelines extended use of the Specialist Rehabilitation Prescription into other areas of clinical practice (including critical care, neurosciences and stroke).[122]

In December 2013, the Royal College of Physicians published National Clinical Guidelines for management of patients in prolonged disorders of consciousness (vegetative and minimally conscious states) following profound brain injury.[219] The guidelines built in tools which were tested and further developed within this programme.[220]

## *Future directions for research*

This programme represents a substantial body of research, which has improved our understanding of the rehabilitation needs of patients with complex disability, the resources that are required to manage them, and the outcomes that may be expected.

It has also provided the Department of Health with critical information about the costs of rehabilitation services, currencies to provide fair reimbursement for cost-efficient intervention, and the scale of cost savings that may be derived from timely rehabilitation interventions.

Although the programme has delivered its key targets so far as it was possible to do so, this remains a time of major change and development within the NHS, and there is still much to be done:

- Development of datasets, tariffs and commissioning currencies for community-based services is still in its infancy – an area that Monitor and NHS England are only just starting to tackle even now.
- Patients with severely challenging behaviour following brain injury (i.e. verbal and physical aggression) continue to present a major conundrum for clinicians and commissioners. Neuro-behavioural rehabilitation services are largely provided through the independent sector and are not yet subject to the same rigorous evaluation and bench-marking as the NHS services. There is still uncertainty about the optimum resources required to manage such patients, and the most appropriate tools for outcome evaluation.
- As the acute care pathways become more adept at saving patients who would otherwise have died at the roadside, more patients are coming through to rehabilitation in prolonged disorders of consciousness. The National Clinical guidelines highlighted a paucity of information both about the numbers of such patients and their long-term prognosis for recovery. They called for a national registry to record progress alongside serial records of responsiveness, using an agreed set of validated tools. The UKROC dataset provides the obvious repository for such information, but will require further development to accommodate this information.

These and other developments will be the subject of a follow-on grant application to continue this important and highly productive applied programme of health services research and development.

## *References*


1. The Neurological Alliance. Neuro numbers, a brief review of the numbers of people in the UK with a neurological condition. London: The Neurological Alliance, 2003.

2. The National Service Framework for Long-term Conditions. London: Department of Health, 2005.

3. Department of Health. Reforming NHS Financial Flows - Introducing payment by results. London: Department of Health: 2002.

4. Turner-Stokes L, McCrone P, Jackson D, Siegert RJ. The Needs and Provision Complexity Scale: a multicentre prospective cohort analysis of met and unmet needs and their cost implications for patients with complex neurological disability. *BMJ Open* 2013;**3: doi 10.1136/bmjopen-2012-002353**

5. Medical rehabilitation in 2011 and beyond. London: Royal College of Physicians, 2011.

6. Turner-Stokes L, Nair A, Disler P, Wade D. Multi-disciplinary rehabilitation for acquired brain injury in adults of working age. CD004170. *The Cochrane Database of Systematic Reviews. Oxford: Update software* 2005 (Update July 2009);**Issue 3**

7. Khan F, Turner-Stokes L, Ng L, Kilpatrick T. Multidisciplinary rehabilitation for adults with multiple sclerosis (Full review). *Cochrane Database of Systematic Reviews* 2007;**Apr 18**(2):CD006036.

8. Turner-Stokes L. Evidence for the effectiveness of multi-disciplinary rehabilitation following acquired brain injury: a synthesis of two systematic approaches. *Journal of Rehabilitation Medicine* 2008 Oct;**40**(9):691-701

9. Turner-Stokes L. The evidence for the cost-effectiveness of rehabilitation following acquired brain injury. *Clinical Medicine* 2004 Jan-Feb;**4**(1):10-12

10. Turner-Stokes L, Paul S, Williams H. Efficiency of specialist rehabilitation in reducing dependency and costs of continuing care for adults with complex acquired brain injuries. *Journal of Neurology, Neurosurgery and Psychiatry* 2006 May;**77**(5):634-639

11. Turner-Stokes L. Cost-efficiency of longer-stay rehabilitation programmes: can they provide value for money? *Brain Injury* 2007;**21**(10):1015-1021

12. Physical disability in 1986 and beyond. London: Royal College of Physicians, 1986.

13. Carter DC. Review of commissioning arrangements for specialised services. The Warner report. London: Department of Health, 2006.

14. National Definition Set for Specialised Services. Third edition. London: Department of Health, 2009.

15. Specialised Services National Definitions Set (3rd edition) No. 7 Specialised Rehabilitation Services for Brain Injury and Complex Disability. London: Department of Health, 2009.

16. Turner-Stokes L, Disler R, Williams H. The Rehabilitation Complexity Scale: a simple, practical tool to identify 'complex specialised' services in neurological rehabilitation. *Clinical Medicine* 2007;**7**(6):593-599

17. Specialist neuro-rehabilitation services: providing for patients with complex rehabilitation needs. London: British Society of Rehabilitation Medicine, 2010.

18. Specialist neuro-rehabilitation services: providing for patients with complex rehabilitation needs. London: British Society of Rehabilitation Medicine, 2015.

19. Braddom RL. Medicare Funding for Inpatient Rehabilitation: How Did We Get to This Point and What Do We Do Now? *Archives of Physical Medicine and Rehabilitation* 2005 2005/7//;**86**(7):1287-1292

20. Zorowitz RD. Inpatient rehabilitation facilities under the prospective payment system: lessons learned. *European Journal of Physical & Rehabilitation Medicine* 2009;**45**(2):259-263

21. Stineman MG, Kallen MA, Thompson C, Gage B. Challenges in paying for effective stays. *Medical Care* 2005;**43**(9):841-843

22. Wikipedia. Case mix 2010 [updated 26/8/2010; cited 2010 27/11/2010]. Available from: http://en.wikipedia.org/wiki/Case_mix.

23. InterRAI. InterRAI netwok 2006 [cited 2010 27/11/2010]. Available from: http://www.interrai.org/section/view/.

24. Fetter R, editor. DRGs their design and development. Ann Arbour: Health Administration Press; 1991.

25. Vladeck BC, Kramer PS. Casemix measures: DRGs and alternatives. *Annual Review of Public Health* 1988;**9**:333-359

26. Schreyögg J, Stargardt T, Velasco-Garrido M, Busse R. Defining the “Health Benefit Bas ket” in nine European countries: Evidence from the European Union Health BASKET Project. *European Journal of Health economics* 2005;**Suppl 1**:6:2–10

27. Reforming NHS financial flows introducing payment by results. London: Department of Health, 2002.

28. Boyle S. Payment by results in England. *Euro Observer* 2005;**7**(4):http://www.euro.who.int/Document/Obs/EuroObserver7_4.pdf

29. International Statistical Classification of Diseases and Related Health Problems. 10th revision. Geneva: World Health Organisation, 2007.

30. Office of Population, Censuses and Surveys Classification of Surgical Operations and Procedures 4th revision. Crown Copyright: Connecting for Health, 2006.

31. Healthcare Resource Group 4: Information Standards Board for Health and Social Care; 2008 [cited 2010 6.6.10]. Available from: http://www.isb.nhs.uk/docs/hrg4/.

32. Turner-Stokes L, Paul S, Williams H. Efficiency of specialist rehabilitation in reducing dependency and costs of continuing care for adults with complex acquired brain injuries.[see comment]. *Journal of Neurology, Neurosurgery & Psychiatry* 2006;**77**(5):634-639

33. Batavia AI, DeJong G. Prospective payment for medical rehabilitation: The DHHS report to congress. *Arch Phys Med Rehabil* 1988;**69**:377

34. Heinemann AW, Linacre JM, Wright BD, Hamilton BB, Granger C. Relationships between impairment and physical disability as measured by the Functional Independence Measure. *Archives of Physical Medicine & Rehabilitation* 1993;**74**(6):566-573

35. Stineman MG, Tassoni CJ, Escarce JJ, Goin JE, Granger CV, Fiedler RC, et al. Development of function-related groups version 2.0: a classification system for medical rehabilitation. *Health Services Research.* 1997;**32**(4):529-548

36. Lee LA, Eagar KM, Smith MC. Subacute and non-acute casemix in Australia. *Medical Journal of Australia* 1998;**169**(5):19

37. Eagar K, Cromwell D, Kennedy C, Lee L. Classifying sub-acute and non-acute patients: results of the New South Wales Casemix Area Network study. *Australian Health Review* 1997;**20**(2):26-42

38. Webster F. Development of a casemix classification system for inpatient rehabilitation services: stage 1 of the Victorian Rehabilitation Project. *Australian Health Review* 1996;**19**(3):81-92

39. Eagar K. The Australian National Sub-Acute and Non-Acute Patient casemix classification. *Australian Health Review* 1999;**22**(3):180-196

40. Nyein K, Thu A, Turner-Stokes L. Complex specialized rehabilitation following severe brain injury: a UK perspective. *Journal of Head Trauma Rehabilitation* 2007;**22**(4):239-247

41. Options for the future of Payment by Results: 2008/09 to 2010/11. London: Department of Health, 2007.

42. Options for the future of Payment by Results: 2008/9 to 2010/11. Department of Health consultation paper. London: Department of Health, June 2007.

43. NIHR. Cost-efficient service provision in neurorehabilitation: defining needs, costs and outcomes for people with Long term neurological conditions (RP-PG-0407-10185). Lead applicant: Professor Lynne Turner-Stokes. Northwick Park Hospital. London: National Institute for Health Research Programme Grant for Applied Research, 2008.

44. Specialist Rehabilitation for Patients with HIghly Complex Needs: D02 Service Specification London2013 [cited 2014]. Available from: [http://www.england.nhs.uk/wp-content/uploads/2014/04/d02-rehab-pat-high-needs-0414.pdf](http://www.england.nhs.uk/wp-content/uploads/2014/04/d02-rehab-pat-high-needs-0414.pdf).

45. Turner-Stokes L, Ward CD. BSRM Standards for Rehabilitation Services, mapped on to the National Service Framework for long-term conditions. London: British Society of Rehabilitation Medicine, 2009.

46. Turner-Stokes L, Williams H, Sephton K, Rose H, Harris S, Thu A. Engaging the hearts and minds of clinicians in outcome measurement –the UK rehabilitation outcomes collaborative approach. *Disabiity and Rehabilitation* 2012

47. Turner-Stokes L, Williams H, Siegert RJ. The Rehabilitation Complexity Scale version 2: a clinimetric evaluation in patients with severe complex neurodisability. *Journal of Neurology, Neurosurgery & Psychiatry* 2010;**81**(2):146-153

48. Turner-Stokes L, Scott H, Williams H, Siegert RJ. The Rehabilitation Complexity Scale – Extended: detection of patients with highly complex needs. *Disability and Rehabilitation* 2012;**34**(9):715-720

49. Turner-Stokes L, Tonge P, Nyein K, Hunter M, Nielson S, Robinson I. The Northwick Park Dependency Score (NPDS): a measure of nursing dependency in rehabilitation. *Clinical Rehabilitation* 1998;**12**(4):304-318

50. Siegert RJ, Turner-Stokes L. Psychometric evaluation of the Northwick Park Dependency Scale. *Journal of Rehabilitation Medicine* 2010;**42**(10):936-943.

51. Turner-Stokes L, Nyein K, Halliwell D. The Northwick Park Care Needs Assessment (NPCNA): a directly costable outcome measure in rehabilitation. *Clinical Rehabilitation* 1999;**13**(3):253-267

52. Turner-Stokes L, Shaw A, Law L, Rose H. Development and initial validation of the Northwick Park Therapy Dependency Assessment. *Clinical Rehabilitation* 2009;**23**(10):922-937

53. Turner-Stokes L. The Northwick Park Therapy Dependency Score (NPTDA): Development, preliminary evaluation and application. Department of Health R&D Project Grant Report ref 030/0066: 2006.

54. Wade DT, Collin C. The Barthel ADL index: a standard measure of physical disability? *International Disability Studies* 1988;**10**:64-67

55. Hamilton BB, Granger CV, Sherwin FS, Zielezny M, Tashman JS. A uniform national data system for medical rehabilitation. In: Fuhrer JM, editor. *Rehabilitation outcomes: analysis and measurement*. Baltimore: Brookes; 1987. p. 137-147.

56. Turner-Stokes L. Standardised outcome assessment in brain injury rehabilitation for younger adults. *Disability and Rehabiliattion* 2002;**24**(7):383-389

57. Skinner A, Turner-Stokes L. The use of standardized outcome measures in rehabilitation centres in the UK. *Clinical Rehabilitation* 2006;**20**(7):609-615

58. Turner-Stokes L. Goal attainment scaling in rehabilitation; a practical guide. *Clinical Rehabilitation* 2009;**23**:362-370

59. Turner-Stokes L, Williams H, J. J. Goal Attainment Scaling: does it provide added value as a person-centred measure for evaluation outcome in neurorehabilitation following acquired brain injury? *Journal of Rehabilitation Medicine* 2009;**41**(7):528-535

60. (Governance Arrangements for Research Ethics Committees (GAfREC) 2011). http://www.hra.nhs.uk/resources/before-you-apply/is-nhs-rec-review-required/.

61. Information Governance Policy. London: NHS England2013. Available from: http://www.england.nhs.uk/wp-content/uploads/2013/06/ig-policy-1.1.pdf.

62. Feinstein AR. Clinimetric perspectives. *Journal of Chronic Disease* 1987;**40**(6):635-640

63. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;**60**(1):34-42

64. Molenaar IW, Sijtsma K, Boer P. MSP5 for Windows: A program for Mokken scale analysis for polytomous items: (Version 5.0) User's Manual iec ProGAMMA; 2000.

65. Mokken RJ, Lewis C. A nonparametric approach to the analysis of dichotomous items. . *Applied Psychological Measurement.* 1982;**6**(4):417-430

66. Sijtsma K, Molenaar IW. Introduction to nonparametric item response theory. In: Jaeger RM, editor. Thousand Oaks, California: Sage Publications; 2002.

67. Van Schuur WH. Mokken scale analysis: Between the Guttman scale and parametric item response theory. In: JH F, editor. *Political Analysis*. 11. Ann Arbor: University of Michigan Press; 2003. p. 139-163.

68. Loevinger J. The technic of homogeneous tests compared with some aspects of "scale analysis" and factor analysis. *Psychological Bulletin* 1948;**45**:507-530

69. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;**33**:159-174

70. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting change in health status. *Medical Care* 1989;**27**:S178-189

71. Cohen J. Statistical power analysis for the behavioural sciences. Second ed. Hillsdale, New Jersey: Lawrence Erlbaum Associates; 1988.

72. Turner-Stokes L, Williams H, Sephton K, Rose H, Harris S, Thu A. Engaging the hearts and minds of clinicians in outcome measurement – the UK Rehabilitation Outcomes Collaborative approach. *Disability and Rehabilitation* 2012;**34**(22):1871-1879

73. Turner-Stokes L, Williams H, Sephton K. Measuring caseload complexity using the extended Rehabilitation Complexity Scale (RCS-E) - a UK multi-centre analysis. American Academy of Physical Medicine and Rehabilitation, Atlanta USA, November 2012 Physical Medicine & Rehabilitation 4(10) S224.

74. Specialist Nursing Home Care for people with Complex Neurological Disability: Guidance to Best Practice. London: British Society of Rehabilitation Medicine, 2013.

75. National Definition Set for Specialised Services No 7: "Complex specialised rehabilitation for brain injury and complex disability (Adult)". Third Edition. London: Department of Health, 2009.

76. Nyein K, Turner-Stokes L, Robinson I. The Northwick Park Care Needs Assessment (NPCNA): A measure of community care needs: Sensitivity to change during rehabilitation. *Clinical Rehabilitation* 1999;**13**(6):482-491

77. Turner-Stokes L, Tonge P, Nyein K, Hunter M, Nielson S, Robinson I. The Northwick Park Dependency Score (NPDS): a measure of nursing dependency in rehabilitation. *Clinical Rehabilitation* 1998;**12**:304-318

78. Svensson S, Sonn U, Sunnerhagen KS. Reliability and validity of the Northwick Park Dependency Score (NPDS) Swedish version 6.0. *Clinical Rehabilitation* 2005 Jun;**19**(4):419-425

79. Post MW, Visser-Meily JM, Gispen LS. Measuring nursing needs of stroke patients in clinical rehabilitation: a comparison of validity and sensitivity to change between the Northwick Park Dependency Score and the Barthel Index. *Clinical Rehabilitation.* 2002;**16**(2):182-189

80. Plantinga E, Tiesinga LJ, van der Schans CP, Middel B. The criterion-related validity of the Northwick Park Dependency Score as a generic nursing dependency instrument for different rehabilitation patient groups. *Clinical Rehabilitation* 2006;**20**(10):921-926

81. Griffiths P, Sironi C. Care needs and point prevalence of post-acute patients in the acute medical wards of an Italian hospital. *International Journal of Nursing Studies* 2005 Jul;**42**(5):507-512

82. Hatfield A, Hunt S, Wade DT. The Northwick Park Dependency Score and its relationship to nursing hours in neurological rehabilitation. *Journal of Rehabilitation Medicine.* 2003;**35**(3):116-120

83. Turner-Stokes L, Nyein K, Turner-Stokes T, Gatehouse C. The UK FIM+FAM: development and evaluation. *Clinical Rehabilitation* 1999;**13**(4):277-287

84. Williams H, Harris R, Turner-Stokes L. Northwick Park Care Needs Assessment: adaptation for inpatient neurological rehabilitation settings. *Journal of Advanced Nursing* 2007;**59**(6):612-622

85. Williams F, Harris R, L. T-S. Work sampling: a quantitative analysis of nursing activity in a neuro-rehabilitation setting. *Journal of Advanced Nursing* (in press)

86. Nyein K, McMichael L, Turner-Stokes L. Can a Barthel score be derived from the FIM? *Clinical Rehabilitation* 1999;**13**(1):56-63

87. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology* 2007;**60**(1):34-42

88. Hatfield A, Hunt S, Wade DT. The Northwick Park dependency score and its relationship to nursing hours in neurological rehabilitation. *Journal of Rehabilitation Medicine* 2003;**35**:116-120

89. Svensson S, Sonn U, Sunnerhagen KS. Reliability and validity of the Northwick Park Dependency Score (NPDS) Swedish version 6.0. *Clinical Rehabilitation* 2005;**19**:419-425

90. Williams H, Harris R, Turner-Stokes L. Northwick Park Care Needs Assessment: adaptation for inpatient neurological rehabilitation settings. *Journal of Advanced Nursing* 2007;**59**(6):612-622

91. Williams F, Harris R, Turner-Stokes L. Work sampling: a quantitative analysis of nursing activity in a neuro-rehabilitation setting. *Journal of Advanced Nursing* (2009);**65**(10):2097-2107

92. Post MWM, Visser-Meily JMA, Gispen L. Measuring nursing needs of stroke patients in clinical rehabilitation: a comparison of validity and sensitivity to change between the Northwick Park Dependency Score and the Barthel Index. *Clinical Rehabilitation* 2002;**16**:182-189

93. Plantinga E, Tiesinga LJ, van der Schans CP, Middel B. The criterion-related validity of the Northwick Park Dependency Score as a generic nursing dependency instrument for different rehabilitation patient groups. *Clinical Rehabilitation* 2006;**20**:921-926

94. Heinemann AW, Linacre JM, Wright BD, Hamilton BB, Granger C. Prediction of rehabilitation outcomes with disability measures. *Archives of Physical Medicine & Rehabilitation* 1994;**75**(2):133-143

95. NHSE. Specialist Rehabilitation for Patients with HIghly Complex Needs:  D02 Service Specification London 2013 [cited 2014]. Available from: http://www.england.nhs.uk/wp-content/uploads/2014/04/d02-rehab-pat-high-needs-0414.pdf.

96. Siegert R, Turner-Stokes L, McCrone P, Playford D, Fleminger S, Bassett P, et al. Evaluation of Community Rehabilitation Service Delivery in Long-Term Neurological Conditions. National Institute for Health Research  Services Delivery and Organisation (SDO) programme; 2008.

97. Siegert RJ, McCrone P, Jackson D, Bassett P, Playford D, Fleminger S, et al. Evaluation of Community Rehabilitation Service Delivery in Long-Term Neurological Conditions. Final report. NIHR Service Delivery and Organisation programme, 2013.

98. Siegert RJ, Jackson DM, Playford ED, Fleminger S, Turner-Stokes L. A longitudinal, multicentre, cohort study of community rehabilitation service delivery in long-term neurological conditions. *BMJ Open* 2014;**4**(e004231.)

99. Williams H, Harris R, Turner-Stokes L. Can the Northwick Park Care Needs Assessment be used to estimate nursing staff requirements in an in-patient rehabilitation setting? *Clinical Rehabilitation* 2007;**21**(6):535-544

100. Williams H, Harris R, Turner-Stokes L. Northwick Park Care Needs Assessment: adaptation for inpatient neurological rehabilitation settings. *Journal of Advanced Nursing* 2007 Sep;**59**(6):612-622

101. Williams H, Harris R, ., Turner-Stokes L. Work sampling: a quantitative analysis of nursing activity in a neuro-rehabilitation setting. *Journal of Advanced Nursing* 2009;**65**(10):2097-2107

102. Kwakkel G. Impact of intensity of practice after stroke: issues for consideration. *Disability & Rehabilitation* 2006;**28**(13-14):823-830

103. Shiel A, Burn JPS, Henry D, Clark J, Wilson BA, Burnett ME, et al. The effects of increased rehabilitation therapy after brain injury: results of a prospective controlled trial. *Clinical Rehabilitation* 2001;**15**(5):501-514

104. Slade A, Chamberlain MA, Tennant A. A randomised controlled trial to determine the effect of intensity of therapy on length of stay in a neurological rehabilitation setting. *Journal of Rehabilitation Medicine* 2002;**34**(6):260-266

105. Gladman J, Barer D, Langhorne P. Specialist rehabilitation after stroke (editorial). *Brit Med J* 1996;**312**:1623-1624

106. Ballinger C, Ashburn A, Low J, Roderick P. Unpacking the black box of therapy -- a pilot study to describe occupational therapy and physiotherapy interventions for people with stroke. *Clinical Rehabilitation* 1999;**13**(4):301-309

107. Dejong G, Horn SD, Gassaway JA, Slavin MD, Dijkers MP. Toward a taxonomy of rehabilitation interventions: Using an inductive approach to examine the "black box" of rehabilitation. *Archives of Physical Medicine & Rehabilitation* 2004;**85**(4):678-686

108. De Wit L, Kamsteegt H, Yadav B, Verheyden G, Feys H, De Weerdt W. Defining the content of individual physiotherapy and occupational therapy sessions for stroke patients in an inpatient rehabilitation setting. Development, validation and inter-rater reliability of a scoring list. *Clinical Rehabilitation* 2007;**21**(5):450-459

109. Bode RK, Heinemann AW, Semik P, Mallinson T. Patterns of therapy activities across length of stay and impairment levels: peering inside the "black box" of inpatient stroke rehabilitation. *Archives of Physical Medicine & Rehabilitation* 2004;**85**(12):1901-1908

110. Latham NK, Jette DU, Slavin M, Richards LG, Procino A, Smout RJ, et al. Physical therapy during stroke rehabilitation for people with different walking abilities. *Archives of Physical Medicine & Rehabilitation* 2005;**86**(12 Suppl 2):S41-S50

111. Richards LG, Latham NK, Jette DU, Rosenberg L, Smout RJ, DeJong G. Characterizing occupational therapy practice in stroke rehabilitation. *Archives of Physical Medicine & Rehabilitation* 2005;**86**(12 Suppl 2):S51-S60

112. van Langeveld SA, Post MW, van Asbeck FW, Postma K, Leenders J, Pons K. Feasibility of a classification system for physical therapy, occupational therapy, and sports therapy interventions for mobility and self-care in spinal cord injury rehabilitation. *Archives of Physical Medicine & Rehabilitation* 2008;**89**(8):1454-1459

113. van Langeveld SA, Post MW, van Asbeck FW, Postma K, Ten Dam D, Pons K. Development of a classification of physical, occupational, and sports therapy interventions to document mobility and self-care in spinal cord injury rehabilitation. *Journal of Neurologic Physical Therapy* 2008;**32**(1):2-7

114. Turner-Stokes L, Shaw A, Law J, Rose H. Development and initial validation of the Northwick Park Therapy Dependency Assessment. *Clinical Rehabilitation* 2009 Oct;**23**(10):922-937

115. Alexandrescu R, Siegert RJ, Turner-Stokes L. The Northwick Park Therapy Dependency Assessment scale: a psychometric analysis from a large multicentre neuro-rehabilitation dataset. *Disability & Rehabilitation* 2015 (E-pub ahead of print. Jan 19:1-8.)

116. Brott T, Adams Jr HP, Olinger CP, Marler JR, Barsan WG, Biller J, et al. Measurements of acute cerebral infarction: a clinical examination scale. *Stroke* 1989;**20**:864-870

117. Maynard Jr. FM, Bracken MB, Creasey G, Ditunno Jr. JF, Donovan WH, Ducker TB, et al. International standards for neurological and functional classification of spinal cord injury. *Spinal Cord* 1997;**35**:266-274

118. Information Governance Policy London: NHS England; available from:2013. Available from: http://www.england.nhs.uk/wp-content/uploads/2013/06/ig-policy-1.1.pdf.

119. International classification of functioning, disability and health. Geneva: World Health Organisation, 2002 ISBN 91 4 154542 9.

120. Turner-Stokes L, Thu A, Williams H, Casey R, Rose H, Siegert RJ. The Neurological Impairment Scale: reliability and validity as a predictor of functional outcome in neurorehabilitation. *Disability and Rehabilitation* 2014;**36**(1):23-31

121. Siegert RJ, Turner-Stokes L, McCrone PM, Jackson DM, Bassett P, Playford ED, et al. Evaluation of Community Rehabilitation Service Delivery in Long-Term Neurological Conditions: Final report. London: National Institute of Heath Research, Health Services and Delivery Research Programme (Grant No: 0001833), 2012.

122. Specialist Rehabilitation in the Trauma pathway: BSRM core standards version 1.3 London: British Society of Rehabilitation Medicine, 2013.

123. Hall KM, Hamilton BB, Gordon WA, Zasler ND. Characteristics and comparisons of functional assessment indices: Disability Rating Scale, Functional Independence Measure, and Functional Assessment Measure. *Journal of Head Trauma Rehabilitation* 1993;**8**(2):60-74

124. Hamilton BB, Laughlin JA, Fiedler RC, Granger CV. Interrater reliability of the 7-level Functional Independence Measure (FIM). *Scandinavian Journal of Rehabilitation Medicine* 1994;**26**(3):115-119

125. Granger CV, Hamilton BB, Linacre JM, Heinemann AW, Wright BD. Performance profiles of the functional independence measure. *American Journal of Physical Medicine & Rehabilitation* 1993;**72**(2):84-89

126. Corrigan JD, Smith-Knapp K, Granger CV. Validity of the functional independence measure for persons with Traumatic Brain Injury. *Archives of Physical Medicine & Rehabilitation* 1997;**78**(8):828-834

127. Ottenbacher KJ, Hsu Y, Granger CV, Fiedler RC. The reliability of the functional independence measure: A quantitative review. *Archives of Physical Medicine & Rehabilitation* 1996;**77**(12):1226-1232

128. Hall KM, Mann N, High WMJ, Wright J, Kreutzer JS, Wood D. Functional measures after traumatic brain injury: Ceiling effects of FIM, FIM+FAM, DRS, and CIQ. *Journal of Head Trauma Rehabilitation* 1996;**11**(5):27-39

129. Wright J. The Center for Outcome Measurement in Brain Injury [17.12.12]. Available from: http://www.tbims.org/combi.

130. Law J, Fielding B, Jackson D, Turner-Stokes L. The UK FIM+FAM Extended Activities of Daily Living module: evaluation of scoring accuracy and reliability. *Disability & Rehabilitation* 2009;**31**(10):825-830

131. Turner-Stokes L, Sutch S, Dredge R. Healthcare tariffs for specialist inpatient neurorehabilitation services: Rationale and development of a UK casemix and costing methodology. *Clinical Rehabilitation* 2011;**26**(3):264-279

132. Turner-Stokes L, Siegert RJ. A comprehensive psychometric evaluation of the UK FIM + FAM. *Disability and Rehabilitation* 2013;**35**(22):1885-1895

133. Linn RT, Blair RS, Granger CV, Harper DW, O'Hara PA, Maciura E. Does the functional assessment measure (FAM) extend the functional independence measure (FIM) instrument? A rasch analysis of stroke inpatients. *Journal of Outcome Measurement* 1999;**3**(4):339-359

134. Tesio L, Cantagallo A. The functional assessment measure (FAM) in closed traumatic brain injury outpatients: a Rasch-based psychometric study. *Journal of Outcome Measurement* 1998;**2**(2):79-96

135. Hawley CA, Taylor R, Hellawell DJ, Pentland B. Use of the functional assessment measure (FIM+FAM) in head injury rehabilitation: a psychometric analysis. *Journal of Neurology, Neurosurgery & Psychiatry* 1999;**67**(6):749-754

136. Hobart JC, Lamping DL, Freeman JA, Langdon DW, McLellan DL, Greenwood RJ, et al. Evidence-based measurement: which disability scale for neurologic rehabilitation? *Neurology* 2001;**57**(4):639-644

137. Donaghy S, Wass PJ. Interrater reliability of the Functional Assessment Measure in a brain injury rehabilitation program. *Archives of Physical Medicine & Rehabilitation* 1998;**79**(10):1231-1236

138. McPherson K, Berry A, Pentland B. Relationships between cognitive impairments and functional performance after brain injury, as measured by the functional assessment measure (FIM + FAM). *Neuropsychological Rehabilitation* 1997;**7**(3):241-257

139. McPherson KM, Pentland B. Disability in patients following traumatic brain injury - Which measure? *International Journal of Rehabilitation Research* 1997;**20**(1):1-10

140. McPherson KM, Pentland B, Cudmore SF, Prescott RJ. An inter-rater reliability study of the Functional Assessment Measure (FIM + FAM). *Disability & Rehabilitation* 1996;**18**(7):341-347

141. Hall KM, Bushnik T, Lakisic-Kazazic B, Wright J, Cantagallo A. Assessing traumatic brain injury outcome measures for long-term follow-up of community-based individuals. *Archives of Physical Medicine & Rehabilitation* 2001;**82**(3):367-374

142. Powell J, Heslin J, Greenwood R. Community based rehabilitation after severe traumatic brain injury: a randomised controlled trial. *Journal of Neurology, Neurosurgery & Psychiatry.* 2002;**72**(2):193-202

143. Seel RT, Wright G, Wallace T, Newman S, Dennis L. The utility of the FIM+FAM for assessing traumatic brain injury day program outcomes. *Journal of Head Trauma Rehabilitation* 2007;**22**(5):267-277

144. Vogele B, Lehmann S, Aschlimann A. Translation, cross-cultural adaption, and reliability of the German version of the Functional Assessment Measure. *Neurologie und Rehabilitation* 2009;**15**:308-314

145. Alcott D, Dixon K, Swann R. The reliability of the items of the functional assessment measure (FAM): Differences in abstractness between FAM items. *Disability & Rehabilitation* 1997;**19**(9):355-358

146. Gurka JA, Felmingham KL, Baguley IJ, Schotte DE, Crooks J, Marosszeky JE. Utility of the functional assessment measure after discharge from inpatient rehabilitation. *Journal of Head Trauma Rehabilitation* 1999;**14**(3):247-256

147. Wilson FC, Wheatley-Smith L, Downes C. Analysis of intensive outpatient neuro-rehabilitation outcomes using FIM+FAM (UK). *NeuroRehabilitation* 2009;**24**(4):377-382

148. Jorge L, Marchi G, Hara P, Battistella L. Brazilian version of the Functional Assessment Measure: cross-cultural adaptation and relaibility evaluation. *International Journal of Rehabilitation Research* 2011;**34**:89-91

149. Tabachnick BG, Fidell LS. Using multivariate statistics (4th edition). Needham Heights, Massachusetts: Allyn & Bacon; 2001.

150. Thompson B. Exploratory and confirmatory factor analysis: Understanding concepts and applications. . Washington, D.C: American Psychological Association, 2004.

151. Rasch G. Probabilistic models for some intelligence and attainment tests. . (Copenhagen, Danish Institute for Educational Research): Chicago: The University of Chicago Press, 1960.

152. Lundgren Nilsson A, Tennant A. Past and present issues in Rasch analysis: the functional independence measure (FIM) revisited. *Journal of Rehabilitation Medicine* 2011;**Oct; 43(10)**:884-891

153. Nayar M, Vanderstay R, Siegert RJ, Turner-Stokes L. The UK Functional Assessment Measure (UK FIM+FAM): a First Formal Psychometric Evaluation in Patients Undergoing Rehabilitation Following a Stroke. 2015

154. Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. *Maryland State Medical Journal* 1965;**14**:61-65

155. Collin C, Wade DT, Davis S, Horne V. The Barthel Index: A reliability study. *International Disability Studies* 1988;**10**:61-63

156. Turner-Stokes L, Williams H, Rose H, Harris S, Jackson D. Deriving a Barthel Index from the Northwick Park Dependency Scale and the Functional Independence Measure: are they equivalent? *Clinical Rehabilitation* 2010 Dec;**24**(12):1121-1126

157. Landis JR, Koch GG. A one-way components of variance model for categorical data. *Biometrics* 1977;**33**:671-679

158. Hurn J, Kneebone I, Cropley M. Goal setting as an outcome measure: A systematic review. *Clinical Rehabilitation* 2006 Sep;**20**(9):756-772

159. Kiresuk T, Sherman R. Goal attainment scaling: a general method of evaluating comprehensive mental health programmes. *Community Mental Health Journal* 1968;**4**:443-453

160. Rockwood K, Stolee P, Fox RA. Use of goal attainment scaling in measuring clinically important change in the frail elderly.[comment]. *Journal of Clinical Epidemiology.* 1993;**46**(10):1113-1118

161. Schlosser RW. Goal attainment scaling as a clinical measurement technique in communication disorders: a critical review. *Journal of Communication Disorders* 2004;**37**(3):217-239

162. Malec JF, Smigielski JS, DePompolo RW. Goal attainment scaling and outcome measurement in postacute brain injury rehabilitation. *Archives of Physical Medicine & Rehabilitation* 1991;**72**(2):138-143

163. Steenbeek D, Ketelaar M, Galama K, Gorter JW. Goal attainment scaling in paediatric rehabilitation: a critical review of the literature. *Dev Med Child Neurol* 2007;**49**(7):550-556

164. Shefler G, Canetti L, Wiseman H. Psychometric properties of goal-attainment scaling in the assessment of mann's time-limited psychotherapy. *Journal of Clinical Psychology.* 2001;**57**(7):971-979

165. Palisano RJ. Validity of goal attainment scaling in infants with motor delays. *Physical Therapy.* 1993;**73**(10):651-658; discussion 658-660

166. Rockwood K, Howlett S, Stadnyk K, Carver D, Powell C, Stolee P. Responsiveness of goal attainment scaling in a randomized controlled trial of comprehensive geriatric assessment. *Journal of Clinical Epidemiology* 2003;**56**(8):736-743

167. Mackay G, Somerville W. Reflections on goal attainment scaling. *Educational Research* 1996;**2**(38):161-171

168. Steenbeek D, Meester-Delver A, Becher JG, Lankhorst GJ. The effect of botulinum toxin type A treatment of the lower extremity on the level of functional abilities in children with cerebral palsy: evaluation with goal attainment scaling. *Clinical Rehabilitation* 2005;**19**(3):274-282

169. Tennant A. Goal attainment scaling: current methodological challenges. *Disability & Rehabilitation* 2007;**29**(20-21):1583-1588

170. Turner-Stokes L, Williams H. Goal attainment scaling: a direct comparison of alternative rating methods. *Clinical Rehabilitation* 2010;**24**(1):66-73

171. Williams RC, Steig RL. Validity and therapeutic efficiency of individual goal attainment procedures in a chronic pain treatment centre. *Clinical Journal of Pain* 1987;**2**:219-228

172. Ashford S, Turner-Stokes L. Goal attainment for spasticity management using botulinum toxin. *Physiotherapy Research International* 2006;**11**(1):24-34

173. Rockwood K, Joyce B, Stolee P. Use of goal attainment scaling in measuring clinically important change in cognitive rehabilitation patients. *Journal of Clinical Epidemiology.* 1997;**50**(5):581-588

174. Rushton PW, Miller WC. Goal attainment scaling in the rehabilitation of patients with lower-extremity amputations: a pilot study. *Archives of Physical Medicine & Rehabilitation.* 2002;**83**(6):771-775

175. Cieza A, Brockow T, Ewert T, Amman E, Kollerits B, Chatterji S, et al. Linking health-status measurements to the international classification of functioning, disability and health. *J Rehabil Med.* 2002;**34**(5):205-210

176. Stucki G, Grimby G. Forward: Applying the ICF in medicine. *Journal of Rehabilitation Medicine* 2004;**36**:5-6

177. Soberg HL, Finset A, Roise O, Bautz-Holter E. Identification and comparison of rehabilitation goals after multiple injuries: an ICF analysis of the patients', physiotherapists' and other allied professionals' reported goals. *Journal of Rehabilitation Medicine* 2008;**40**(5):340-346

178. Gordon JE, Powell C, Rockwood K. Goal attainment scaling as a measure of clinically important change in nursing-home patients. *Age & Ageing.* 1999;**28**(3):275-281

179. Rauch A, Cieza A, Stucki G. How to apply the International Classification of Functioning, Disability and Health (ICF) for rehabilitation management in clinical practice. *European Journal of Physical & Rehabilitation Medicine* 2008;**44**(3):329-342

180. Kiresuk TJ, Chaote RO. Applications of goal attainment scaling. In: Kiresuk TJ CJ, and Smith A., editor. *Goal Attainment Scaling: Applications, Theory and Measurement*. New Jersey: Lawrence Erlbaum Associates Inc; 1994. p. 61-104.

181. Malec J. The Mayo-Portland Adaptability Inventory The Center for Outcome Measurement in Brain Injury, http://www.tbims.org/combi/mpai (accessed June 14, 2015), 2005.

182. Alderman N, Wood RL, Williams C. The development of the St Andrew's-Swansea Neurobehavioural Outcome Scale: validity and reliability of a new measure of neurobehavioural disability and social handicap. *Brain injury [Validation Studies]* 2011;**25**(1):83-100

183. Wing JK, Curtis RH, Beevor AS. Health of the Nation Outcome Scales (HoNOS). *British Journal of Psychiatry* 1999;**174**(5):432-434

184. Turner-Stokes L, Siegert RJ. The Needs and Provision Complexity Scale: a first psychometric analysis. *In submission, Disabiity and Rehabilitation* 2012

185. Turner-Stokes L, McCrone P, Jackson DM, Siegert RJ. The Needs and Provision Complexity Scale: a multicentre prospective cohort analysis of met and unmet needs and their cost implications for patients with complex neurological disability. *BMJ open* 2013;**3**(2)

186. Turner-Stokes L, Sutch S, Dredge R. International casemix and costing models:lessons for rehabilitation in the UK. *Clinical Rehabilitation* 2011;**26**(3):195-208

187. Sutherland JM, Walker J. Challenges of rehabilitation case mix measurement in Ontario hospitals. *Health Policy* 2008;**85**(3):336-348

188. Heinemann AW, Hamilton B, Linacre JM, Wright BD, Granger C. Functional status and therapeutic intensity during inpatient rehabilitation. *American Journal of Physical Medicine & Rehabilitation* 1995;**74**(4):315-326

189. Heinemann AW, Kirk P, Hastie BA, Semik P, Hamilton BB, Linacre JM, et al. Relationships between disability measures and nursing effort during medical rehabilitation for patients with traumatic brain and spinal cord injury. *Archives of Physical Medicine & Rehabilitation* 1997;**78**(2):143-149

190. Disler PB, Roy CW, Smith BP. Predicting hours of care needed. *Archives of Physical Medicine and Rehabilitation* 1993;**74**(2): 139-143

191. Hoffman JM, Doctor JN, Chan L, Whyte J, Jha A, Dikmen S. Potential impact of the new Medicare Prospective Payment System on reimbursement for traumatic brain injury inpatient rehabilitation. *Archives of Physical Medicine & Rehabilitation.* 2003;**84**(8):1165-1172

192. Evans RL, Halar EM, Hendricks RD, Lawrence KV, Kirk C, Bishop DS. Effects of prospective payment financing on rehabilitation outcome. *International Journal of Rehabilitation Research* 1990;**1**(1):27-36

193. Green J, Gordon R. The development of version 2 of the AN-SNAP casemix classification system. *Australian Health Review* 2007;**31 (suppl 1)**:s68-s78

194. Fries BE, Schneider DP, Foley WJ, Gavazzi M, Burke R, Cornelius E. Refining a case-mix measure for nursing homes: Resource Utilization Groups (RUG-III). *Medical Care* 1994;**32**(7):668-685

195. Arling G, Kane RL, Mueller C, Lewis T. Explaining direct care resource use of nursing home residents: findings from time studies in four states. *Health Services Research* 2007;**42**(2):827-846

196. Hirdes JP, Botz CA, Kozak J, Lepp V. Identifying an appropriate case mix measure for chronic care: evidence from an Ontario pilot study. *Healthcare Management Forum* 1996;**9**(1):40-46

197. Björkgren MA, Häkkinen U, Finne-Soveri UH, Fries BE. Validity and reliability of Resource Utilization Groups (RUG-III) in Finnish long-term care facilities. *Scandinavian Journal of Public Health* 1999;**27**:228 - 234

198. Carpenter GI, Main A, Turner GF. Casemix for the elderly inpatient: Resource Utilization Groups (RUGs) validation project. Casemix for the Elderly Inpatient Working Group. *Age & Ageing* 1995;**24**(1):5-13

199. Hall KM, Mann N, High WMJ, Wright J, Kreutzer JS, Wood D. Functional measures after traumatic brain injury: Ceiling effects of FIM, FIM+FAM, DRS, and CIQ. *Journal of Head Trauma Rehabilitation* 1996;**11**(5):27-39

200. Guide for the Uniform Data Set for Medical Rehabilitation (including the FIM(TM) instrument), Version 5.1. Buffalo, NY Buffalo, NY 14214- 3007: State University of New York at Buffalo, 1997.

201. Turner-Stokes L, Bill A, Dredge R. A cost analysis of specialist inpatient neurorehabilitation services in the UK. *Clinical Rehabilitation* 2011;**26**(3):256-263

202. NHS England. NHS Standard Service specification D02: Specialist Rehabilitation for patients with highly complex needs (All ages) London2013. Available from: [http://www.england.nhs.uk/commissioning/wp-content/uploads/sites/12/2014/04/d02-rehab-pat-high-needs-0414.pdf](http://www.england.nhs.uk/commissioning/wp-content/uploads/sites/12/2014/04/d02-rehab-pat-high-needs-0414.pdf).

203. Acute health clinical costing standards. London: Department of Health, 2009.

204. National Definition Set for Specialised Services No 7: "Complex specialised rehabilitation for brain injury and complex disability (Adult)". London: Department of Health, 2002.

205. Kolehamainen-Aiken RL. Decentralization and human resources: implications and impact. *Human Resources for Health Development* 1997;**2**(1):1-14

206. Saltman RB, Von Otter C. Implementing Planned Markets in Health Care: Balancing Social and Economic Responsibility. . Buckingham: Open University Press 1995.

207. Curtis L. Unit costs of health & social care 2013. Personal Social Services Research Unit, University of Kent Canterbury, 2013.

208. Granger CV, Cotter AC, Hamilton BB, Fiedler RC, Hens MM. Functional assessment scales: a study of persons with multiple sclerosis. *Archives of Physical Medicine & Rehabilitation* 1990;**71**(11):870-875

209. Granger CV, Cotter AC, Hamilton BB, Fiedler RC. Functional assessment scales: A study of persons after stroke. *Archives of Physical Medicine & Rehabilitation* 1993;**74**(2):133-138

210. Chae J, Zorowitz RD, Johnston MV. Functional outcome of hemorrhagic and nonhemorrhagic stroke patients after in-patient rehabilitation. *American Journal of Physical Medicine & Rehabilitation.* 1996;**75**(3):177-182

211. Dombovy ML, Drew-Cates J, Serdans R. Recovery and rehabilitation following subarachnoid haemorrhage. Part I: Outcome after inpatient rehabilitation. *Brain Injury.* 1998;**12**(6):443-454

212. Dombovy ML, Drew-Cates J, Serdans R. Recovery and rehabilitation following subarachnoid haemorrhage: Part II. Long-term follow-up. *Brain Injury* 1998;**12**(10):887-894

213. Streppel KR, Van Harten WH. The Functional Independence Measure used in a Dutch rehabilitating stroke population; a pilot study to assess progress. *International Journal of Rehabilitation Research.* 2002;**25**(2):87-91

214. van der Putten JJ, Hobart JC, Freeman JA, Thompson AJ. Measuring change in disability after inpatient rehabilitation: comparison of the responsiveness of the Barthel index and the Functional Independence Measure. *Journal of Neurology, Neurosurgery & Psychiatry.* 1999;**66**(4):480-484

215. Siegert RJ, Jackson DM, Playford ED, Fleminger S, Turner-Stokes L. A longitudinal, multicentre, cohort study of community rehabilitation service delivery in long-term neurological conditions. *BMJ open* 2014;**4**(2):e004231

216. Turner-Stokes L, Poppleton R, Williams H, Schoewenaars K, Badwan D. Using the UKROC dataset to make the case for resources to improve cost-efficiency in neurological rehabilitation. *Disability and Rehabilitation* 2012;**34**(22):1900-1906

217. Poppleton R, Turner-Stokes L, Schoewenaars K. The journey in trying to achieve eligibility for level 1 status in a specialised rehabilitation service. *Social Care and Neurodisability* 2012;**3**(3):131-139

218. BSRM. Specialist Rehabilitation in the Trauma pathway:  BSRM core standards London: British Society of Rehabilitation Medicine, 2013.

219. Prolonged disorders of consciousness: national clinical guidelines. London: Royal College of Physicians, 2013.

220. Turner-Stokes L, Bassett P, Rose H, Ashford S, Thu A. Serial measurement of Wessex Head Injury Matrix in the diagnosis of patients in vegetative and minimally conscious states: a cohort analysis. *BMJ open* 2015;**5**(4):e006051.