The efficacy of inpatient rehabilitation in multiple sclerosis

Thesis submitted for the degree of Doctor of Philosophy
in the Faculty of Science of the University of London
March 1997

Jennifer Anne Freeman

Department of Clinical Neurology
Institute of Neurology, Queen Square, London
I would like to thank the people with multiple sclerosis who took part in these studies, and the staff at the Neurorehabilitation Unit of the National Hospital for Neurology and Neurosurgery; without their co-operation this thesis would not have been possible.

I am indebted to my supervisors Dr Alan Thompson and Dr Dawn Langdon for their encouragement, tireless patience, and the considerable time and effort that they have given in guiding, advising and motivating me throughout the four years of this thesis. Their support and assistance has been invaluable.

I would also like to thank Dr Jeremy Hobart, for undertaking all of the neurological assessments in these studies, and for providing advice, valuable comments……and entertainment.

To my husband Robert, I am ever grateful for his support, his encouragement, his good humour and friendship.
Inpatient rehabilitation is advocated as an important intervention in the symptomatic management of MS. It is however costly, its effectiveness has not been established, and little is known about the long term carryover of benefit. The basis of this thesis is the design of two closely linking studies which evaluate the effectiveness of inpatient rehabilitation of patients in the progressive phase of MS.

A stratified randomised wait-list controlled trial was undertaken. Sixty six patients were assessed at zero and six weeks with validated measures of impairment, disability, handicap, quality of life and emotional well-being. At baseline both groups were comparable in all variables. At the end of six weeks, although the level of impairment in both groups remained the same, those who participated in rehabilitation (average of 25 days) significantly improved their level of disability (p<0.001) and handicap (p<0.01) compared to those in the wait-list control group.

A longitudinal study investigated the duration of benefit of change. Fifty consecutive patients were assessed, using the same outcome measures as the first trial, on admission, discharge, and three monthly intervals for one year. Twelve month data was collected for 92% of patients. Summary measures were calculated to determine the length of time taken for each individual to return to their baseline level. Over the study period, neurological status declined. Improvements were maintained in disability and handicap for six months,
emotional well-being for seven months, and quality of life (physical component) for ten months.

The conclusions are that despite unchanging impairment, inpatient rehabilitation results in reduced disability and handicap in patients with progressive MS. These benefits are maintained, in part, following discharge into the community, despite worsening neurological status. Carryover of benefits does, however, decline over time.
Publications relevant to this thesis


Table of Contents

Title page ......................................................................................................................... 1

Acknowledgements ........................................................................................................... 2

Abstract ............................................................................................................................ 3

Publications relevant to this thesis ................................................................................. 5

Table of contents .............................................................................................................. 6

List of figures and tables .................................................................................................. 13

Chapter One: Introduction ............................................................................................... 18

1.1 The disease and its consequences ............................................................................. 18

1.1.1 Epidemiology ........................................................................................................ 18

1.1.2 The disease process .............................................................................................. 19

1.1.3 Impact of multiple sclerosis on the individual ...................................................... 21

1.2 The economic cost ..................................................................................................... 22

1.3 Health care management ........................................................................................... 24

1.3.1 Inpatient neurorehabilitation ............................................................................. 25

1.4 The need for evaluation of health care services ....................................................... 29

Chapter Two: Measuring the effectiveness of inpatient rehabilitation:

what outcomes to measure ............................................................................................... 31

2.1 The ICIDH: a framework for measuring the outcome of inpatient rehabilitation ....... 32
2.1.1 The conceptual framework ......................................................32
2.1.2 Terminology .............................................................................33
2.1.3 Application to clinical practice .................................................34
2.1.4 Application of the ICIDH to outcome studies in rehabilitation.36

2.2 Using appropriate measurement instruments ........................................37

2.3 Measuring impairment ........................................................................40
  2.3.1 Instruments for measuring impairment in MS .........................42
  2.3.2 Expanded Disability Status Scale and Functional Systems .......43
  2.3.3 Psychometric properties of the EDSS and FS ..........................47

2.4 Measuring disability ............................................................................55
  2.4.1 Studies in Multiple Sclerosis ...................................................56
  2.4.2 Instruments for measuring disability ........................................58
  2.4.3 The Functional Independence Measure .................................59
  2.4.4 Psychometric properties of the FIM ........................................60

2.5 Measuring handicap ...........................................................................69
  2.5.1 Studies in multiple sclerosis ....................................................71
  2.5.2 Instruments for measuring handicap ........................................74
  2.5.3 The London Handicap Scale ...................................................75
  2.5.4 Psychometric properties of the LHS ........................................77

2.6 Measuring Health-Related Quality of Life ........................................82
  2.6.1 Defining health-related quality of life ......................................83
  2.6.2 Conceptual issues .................................................................84
Chapter Three: Outcome studies of comprehensive MS inpatient rehabilitation

3.1 Early outcome studies ................................................................. 113

3.2 Current methodological dilemmas .............................................. 125

3.3 Rationale for the studies undertaken in this thesis .................... 131

Chapter Four: Experiment One - Randomised Controlled Study .......... 133

4.1 Aim of the study ......................................................................... 133

4.2 Null hypothesis ......................................................................... 133

4.3 Study design ............................................................................. 133
4.4 Subjects and methods ................................................................. 133

4.4.1 Sample size ............................................................................ 133

4.4.2 Recruitment ........................................................................... 134

4.4.3 Procedures ............................................................................. 135

4.4.4 Stratification and randomisation .......................................... 135

4.4.5 Test battery and methods of administration ...................... 137

4.4.6 The rehabilitation service and programme ....................... 141

4.4.7 Quality control of data input ................................................. 144

4.4.8 Statistical analysis ............................................................... 144

4.5 Results ..................................................................................... 146

4.5.1 Sample: baseline characteristics ....................................... 148

4.5.2 The rehabilitation intervention ......................................... 151

4.5.3 Outcome measurement change scores .............................. 153

4.5.4 Self-reported relapses ......................................................... 165

4.6 Discussion ............................................................................... 165

4.6.1 Findings ............................................................................... 165

4.6.2 Evaluation of methodology ............................................... 176

4.7 Conclusion .............................................................................. 184

Chapter Five: Experiment Two - Longitudinal Trial ...................... 185

5.1 Aim of the study ..................................................................... 185

5.2. Study design ......................................................................... 185
5.3 Subjects and methods ................................................................. 185
  5.3.1 Sample size ........................................................................... 185
  5.3.2 Recruitment ......................................................................... 186
  5.3.3 Procedures ........................................................................... 186
  5.3.4 Test battery and methods of administration ....................... 187
  5.3.5 The rehabilitation service and programme........................ 188
  5.3.6 Quality control of data input ............................................... 188
  5.3.7 Statistical analysis .............................................................. 189

5.4 Results ..................................................................................... 193
  5.4.1 Sample baseline characteristics ........................................... 195
  5.4.2 The rehabilitation intervention .......................................... 201
  5.4.3 Self-reported relapses ......................................................... 205
  5.4.4 Re-hospitalisations ............................................................. 205
  5.4.5 Outcome and summary measure scores ............................... 205
  5.4.6 Correlation analyses ........................................................... 224

5.5 Discussion ................................................................................ 228
  5.5.1 Findings .............................................................................. 228
  5.5.2 Outcome and summary measures ....................................... 229
  5.5.3 Correlation analyses ............................................................ 245
  5.5.4 General comments ............................................................. 247
  5.5.5 Evaluation of methodology ............................................... 249

5.6 Conclusion .............................................................................. 253
Chapter Six: Conclusions ..............................................................................................................255

6.1 Tackling the methodological dilemmas encountered .........................................................256
6.2 The methodological challenges that remain ........................................................................258
6.3 The findings of these studies ..............................................................................................261
6.4 Clinical implications ...........................................................................................................262
6.5 The future ..........................................................................................................................267
6.6 Summary ............................................................................................................................277

References ..................................................................................................................................279

Appendices ..................................................................................................................................315

Appendix 1 Items and scoring system of the Functional Systems Scale .................................315
Appendix 2 Items and scoring system of the Expanded Disability Status Scale .......................316
Appendix 3 Items and scoring system of the Functional Independence Measure ....................317
Appendix 4 Items of the London Handicap Scale ....................................................................318-319
Appendix 5 Items of the Short Form Health Survey Questionnaire ........................................320-321
Appendix 6 Items of the General Health Questionnaire .........................................................322
Appendix 7 Information leaflet for patients in the randomised controlled trial ........................323
Appendix 8 Form for demographic and diagnostic data on admission ....................................324
Appendix 9  Intrarater reliability in scoring the FIM..................................325
Appendix 10  Validity of the self-report method for scoring the FIM.......326
Appendix 11  Information leaflet for patients in the longitudinal trial......327
Appendix 12  Form for demographic data on discharge.........................328
Appendix 13  Form for demographic data on follow-up..........................329
Appendix 14  Profile of patients who experienced a rapid reduction in
function between two consecutive assessments.........................330
Figures and Tables

Table 1.1: Average yearly cost estimates of MS in three countries..............23

Figure 1.1: Health care management in people with MS..........................25

Table 1.2: Services central to a rehabilitation service..............................27

Table 2.1: Range of impairments experienced by people with MS..............40

Table 2.2: Instruments for rating impairment in MS.................................42

Table 2.3: Studies evaluating EDSS interrater reliability...........................49

Table 2.4: Studies evaluating FS interrater reliability...............................50

Table 2.5: Range of disabilities and their estimated prevalence in the UK..55

Table 2.6: Instruments used to evaluate disability in MS..........................58

Table 2.7: Interrater reliability studies of the FIM...................................62

Table 2.8: Dimensions of the London Handicap Scale.............................76

Table 2.9: Test retest reliability studies on the LHS.................................78

Table 2.10: Associations between LHS and other health status instruments.80

Figure 2.1: A conceptual and operational model for measuring outcome.....82

Table 2.11: Conceptual domains of the SF-36 ........................................93
Table 2.12: Associations between the SF-36 and other health status instruments.................................................................95

Table 2.13: Contents of the GHQ-28................................................................................................................103

Table 2.14: Associations between sub-scales on the GHQ-28 and clinical ratings..............................................................108

Table 2.15: Validity studies of the GHQ-28..............................................................................................................109

Table 2.16: Summary of the outcomes assessed in this thesis.........................111

Table 3.1: Outcome studies of MS inpatient rehabilitation............. 120-122

Figure 3.1: Patterns of disease activity in MS in most patients.................126

Figure 4.1: Outline of study with sequence of assessment events..............136

Table 4.1: Drug categorisation system..........................................................137

Figure 4.2: Schematic representation of the assessment, selection and rehabilitation service provided by the NRU, NHNN.........142

Table 4.2: Patients failing to meet the study's entry criteria.............146

Table 4.3: Baseline demographic and diagnostic criteria.....................148

Table 4.4: Baseline outcome measurement values...............................149

Table 4.5: Broad areas addressed in the rehabilitation programme...........152

Table 4.6: Drug management of study patients.................................153
Table 4.7: EDSS and FS median change scores, effect size and confidence intervals for treatment and control groups

Table 4.8: FIM change scores for treatment and control groups

Figure 4.3: Boxplot showing the difference in disability change scores between the treatment and control groups

Table 4.9: LHS change scores for treatment and control groups

Figure 4.4: Boxplot chart showing the difference in handicap change scores between the treatment and control groups

Table 4.10: Profile of SF-36 change scores for treatment & control groups

Table 4.11: Effect sizes and 95% confidence intervals for the SF-36 change scores

Table 4.12: GHQ change scores for treatment and control groups

Table 5.1: Patients failing to meet the study’s entry criteria

Table 5.2: Baseline disease characteristics

Table 5.3: Hospital and community services on admission

Table 5.4: Baseline outcome measurement values

Table 5.5: Broad areas addressed in the rehabilitation programme

Table 5.6: Additional hospital and community services recommended at discharge

Table 5.7: Drug management of study patients
Table 5.8: Re-hospitalisations following discharge into the community.....205

Figure 5.1: Group median EDSS scores at each assessment point..........206

Figure 5.2: Distribution of group EDSS scores on admission & 12 months.207

Table 5.9: Group Functional Systems scores at each assessment point...208

Figure 5.3: Individual plots of EDSS scores against time......................209

Figure 5.4: Group median FIM motor scores at each assessment point.....210

Figure 5.5: Individual plots of FIM scores against time........................211

Figure 5.6: Group mean LHS scores at each assessment point..............213

Figure 5.7: Individual plots of LHS scores against time......................214

Figure 5.8a: Group mean SF-36 Physical Health summary scores at each
assessment point.............................................................................216

Figure 5.8b: Group mean SF-36 Mental Health summary scores at each
assessment point.............................................................................216

Figure 5.9a: Individual plots of SF-36 Physical Health Summary scores against
time.....................................................................................................217

Figure 5.9b: Individual plots of SF-36 Mental Health Summary scores against
time.....................................................................................................218

Table 5.10: SF-36 Physical and Mental Health Summary change scores....219

Figure 5.10: Group median GHQ scores at each assessment point.........220
Figure 5.11: Individual plots of GHQ scores against time.............................221

Table 5.11: Summary of data for all outcome measures assessed in the longitudinal study..................................................................................223

Table 5.12: Correlation matrix of baseline outcome measurement scores...225

Table 5.13: Numbers of individual patients progressing from one EDSS category to another during a 12-month period............................232

Table 5.14a: The progression of EDSS scores over the 12 month study period of the longitudinal study.................................................................231

Table 5.14b: The progression of EDSS scores over a 12 month period as determined by Quick and Schapiro's survey...............................233
Multiple Sclerosis (MS) is a chronic ideopathic inflammatory demyelinating disease of the central nervous system (CNS). It usually begins in early adult life, and is said to be the most common neurological cause of severe disability in young people (Barnes 1993; Thompson 1996a).

1.1 The disease and its consequences

1.1.1 Epidemiology

The prevalence of MS varies throughout the world and appears to be influenced by a complex combination of geographical position and genetic background (Poser 1994). It affects approximately 85,000 people in the United Kingdom (UK) and up to three million people world-wide (Thompson and McDonald 1996). On average these rates are higher than in the rest of Europe (Rosati 1994), with parts of the UK having the highest incidence, prevalence and mortality rates in the world (Swingler and Compston 1986). The prevalence rate in the UK is estimated between 95/100,000 (Roberts et al 1991) and 144/100,000 (Phadke and Downie 1987), with an estimated incidence ranging between 4.7 new cases/100,000/year (Roberts et al 1991) to 7.2 new cases/100,000/year (Phadke and Downie 1987). These rates are usually regarded as conservative estimates. This is because of changes in definition and classification, diagnostic problems, levels of health care and the
standard of record keeping (Compston 1990). Almost twice as many females are affected than males with a male:female ratio varying from 1:1.5 to 1:2 (Mathews 1991).

1.1.2 The disease process

The diffuse demyelination, inflammation and gliosis which occurs in MS produces abnormalities in conduction at multiple sites both within the CNS and at the spinal root entry zones of the peripheral nervous system (Thomas et al 1987). This results in a wide variety of clinical signs and symptoms which generally correspond to the area of CNS affected, and may include: cognitive impairment, bladder and bowel dysfunction, sexual dysfunction, dysarthria, dysphagia, visual deficits, sensory disturbance, fatigue, weakness, spasticity, ataxia and reduced mobility (Mathews 1991). In the early stages of MS, clinical recovery from individual episodes of inflammation and demyelination may occur (Sibley 1993). However, with repeated episodes and accumulation of damage, the ability of the axons to remyelinate, adapt and compensate for these degenerative changes is reduced. Eventually persistent and widespread conduction block may occur, resulting in permanent disability (Thompson and McDonald 1992).

Characteristically the course of the disease is extremely variable, unpredictable and generally progressive in nature with no clear-cut or typical sequence of events (Mathews 1991). Classically however it begins with a series of relapses and remissions. In time the remissions tend to be less complete with approximately two thirds of people passing into a progressive phase (secondary progressive) where a gradual accumulation of irreversible, and
often complex and fluctuating, disabilities occur. Up to one third of people do not develop progressive disability and remain relatively unimpaired many years after the onset of the disease (benign MS). A smaller number of people (10-15%) develop progressive disability from onset without relapses and remissions (primary progressive MS) (Thompson and McDonald 1992; Weinshenker 1994).

A wide spectrum of disease activity in terms of the frequency and severity of attacks, the degree of recovery and the development of progressive disability exists in MS (Weinshenker 1994). At one extreme people experience only mild symptoms which barely interfere with function even after 15 years or more of illness, whereas others experience a rapidly deteriorating disease course, with severe impairment after only a few years. The great majority of people fall between these extremes and gradually accumulate disability with increasing time.

Because of the unpredictable nature of MS, attempts to predict outcome in terms of disability have proven extremely difficult (Runmarker et al 1994; Weinshenker et al 1996a), resulting in a large number of clinical studies producing wide ranging and sometimes conflicting results. Confavreaux and colleagues (1980) estimated that moderate disability occurred on average 3.4 years from onset, with severe disability by 9.5 years on average. Weinshenker and Ebers (1987) calculated that 50% of people required walking aids, or use of a wheelchair, within 15 years of diagnosis. Others suggest that the best indicator of prognosis is the functional status of the patient at five or ten years after onset (Kurtzke et al 1977; Thompson and McDonald 1996). Despite these
differences, it is generally agreed that the onset of the secondary progressive phase is the main determinant of disability, with patients whose disease is progressive from onset having the worst long term prognosis (Thompson and McDonald 1992; Runmarker et al 1994; Weinshenker et al 1996a). More recently it has been shown that long-term progression (time taken to require the use of walking aids) and short term progression (worsening over one to three years) differ, and are only weakly correlated to one another (Weinshenker et al 1996a).

MS is rarely a direct cause of death and lifespan is not substantially altered (Weinshenker 1994). The mean duration of disease from onset of symptoms ranges from 25 years (Swingler and Compston 1986) to in excess of 35 years (Poser et al 1989).

1.1.3 Impact of Multiple Sclerosis on the individual

Because MS begins in early adult life and has little effect on longevity, it may have a major long term impact on family, work and social life. Numerous accounts are available describing this (e.g. McLellan et al 1989; Perry 1994). Compared to the general population people with MS are less likely to be engaged in activities of their own choosing such as work, school or homemaking; they receive more inpatient and outpatient hospital care; and they need more personal assistance (Kalb 1995). Patients face not only the prospect and reality of increasing disability, but also the uncertainty of when new relapses will occur or when established disability will set in (Campion 1996). The multiplicity of symptoms means that the physical and psychosocial
problems experienced are wide ranging, variable and often complex. They may evolve over several decades (Thompson 1996b). The needs of the person with MS are therefore many and ever-changing (Campion 1996). They extend from the core medical parameters to include every facet of individual, family and community existence (Perry 1994).

1.2 The economic cost

The wide ranging consequences of MS may result in a major burden of suffering for patients and their families and make substantial demands upon health, social and voluntary sectors throughout the UK (British Society of Rehabilitation Medicine {BSRM} Working Party Report on MS 1993). The socio-economic implications are therefore significant. Estimations of the economic cost to the individual with MS, their families, and society in general, vary widely (Harvey 1995). This may be due not only to the differing costs and prioritisation of resources between geographical areas, but also to methodological differences between studies (e.g. study design, data sources, estimation methodologies, and types of costs included in the calculation).

There is a consensus that while variations in estimations occur (as illustrated in Table 1.1), the costs to society (both direct and indirect) increase as the level of neurological dysfunction increases (Prouse et al 1991; Bourdette et al 1993; Harvey 1995; Holmes et al 1995). For example, Holmes and colleagues (1995) estimated the total cost for each individual to vary from £2,643 for "highly mobile sufferers" to £10,756 for wheelchair users. While these studies do not provide a detailed picture of the economic costs of MS, they begin to give us an
idea of the heavy financial burden that this disease places on both the individual and society.

Table 1.1 Average yearly cost estimates of MS in three countries

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevalence rate</td>
<td>≈ 58/100000</td>
<td>≈ 96/100000</td>
<td>≈ 80/100000</td>
</tr>
<tr>
<td></td>
<td>i.e. 123 000</td>
<td>i.e. 8000</td>
<td>i.e. 87 873</td>
</tr>
<tr>
<td>Total societal costs / annum</td>
<td>Total = £1.05 billion</td>
<td>Direct costs = £4.4 million Indirect costs = £108.5 million</td>
<td>Direct NHS costs = £153 million</td>
</tr>
<tr>
<td>Estimated direct costs per capita / annum</td>
<td>£ 8,512</td>
<td>£ 19,250</td>
<td>£15,000</td>
</tr>
</tbody>
</table>

- Costs have been converted to £ Stirling for comparison purposes.
- The design for all studies was the "cost-of-illness method" based on prevalence rates.
- Estimated costs include both the direct (e.g. drugs, medical services, home care services) and indirect costs (e.g. lost earnings) of MS.
- Estimated costs are referenced from: Great Britain (Holmes et al 1995); USA (Harvey 1995); Sweden (Jonsson 1995).

More specifically, areas such as unemployment can be investigated to estimate the cost of this disease. MS primarily strikes young adults in the midst of their most productive years (mean age of onset is 30 years) and therefore has a dramatic impact on this area, with unemployment rates as high as 75-80% (McLellan et al 1989; Harvey 1995). Many people leave work during the first five years of their illness (Enteen 1995). It has been estimated that in Great Britain, the yearly cost from net earnings foregone, of people with MS and their carers, amount to £395 million with the cost to industry from days lost from work to be £41 million (Holmes et al 1995). This is similar to the experience of other countries (Bourdette et al 1993; Jonsson 1995; Harvey 1995).
In terms of health care costs, it can be seen that Great Britain spends an estimated £153 million per annum on National Health Care Services (Table 1.1). The estimated annual NHS costs range from at least £336 - £4275 per patient, depending upon their level of mobility (Holmes et al 1995). These costs were principally associated with hospital in-patient and outpatient visits, and other non-general practitioner treatment. Drug costs were considered insignificant (Holmes et al 1995). The recent approved use of interferon beta 1-b (at an estimated cost of $US10,000 per person per year for the drug alone, Reingold 1996), and the expected development of other similar drugs is likely to have an escalating effect on these costs.

1.3 Health care management

Strategies to combat the problems of MS encompass a number of different approaches including: its prevention and cure, symptomatic management, and the provision of supportive services. There is no cure for MS. Recent advances in drug therapies which reduce relapse rate offer renewed, if limited hope (e.g. interferon beta 1-b, copolymer 1, interferon beta 1-a). However a clinically meaningful effect on disability has not yet been adequately demonstrated (Thompson and Noseworthy 1996). Progressive disability therefore remains the characteristic experience of the majority of people with MS and symptomatic management and the provision of supportive services continue to be the cornerstone of health care intervention.

Symptomatic management is entirely dependent upon the clinical presentation of each individual and, ideally, changes according to their needs at different points in time. Because of the multiplicity of symptoms and the manner in which
they interact, attempting to treat one symptom without taking account of the others is likely to be unsuccessful (Thompson and McDonald 1996).

Management is frequently spread across the hospital, primary, community care, and voluntary sectors (Figure 1.1). It may include medical or pharmacological measures, surgical intervention, psychological, and physical or rehabilitative methodologies (Scheinberg 1994). Effective management requires the knowledge, expertise, and collaboration of a variety of health care and social service professionals both within and between these different sectors (Langton Hewer 1980; Barnes 1993).

![Image: Figure 1.1 Health care management in people with MS]

**Figure 1.1 Health care management in people with MS**

**1.3.1 Inpatient neurorehabilitation**

Inpatient neurorehabilitation is one such service which focuses on providing a co-ordinated multidisciplinary approach to managing the everyday problems of people with MS. It is recognised as an integral part of health care for disabled people (Neurological Provision-key areas and targets 1991; Association of British Neurologists {ABN} Working Party Report 1992; National Audit Office 1992) and is advocated as an important and effective intervention in MS (BSRM Working Party Report on MS 1993).
The process of inpatient rehabilitation has been described by a number of authors (e.g. Langton-Hewer 1980; Scheinberg et al 1981; Erickson et al 1989; Schapiro 1990; La Rocca and Kalb 1992; Barnes 1993; Mertin 1994; Thompson et al 1994). While no two centres appear to practice or deliver care in an identical way, the literature demonstrates that there is a shared conceptual and practical framework to its clinical practice. This approach is based on a model of comprehensive care (La Rocca et al 1994) which considers that rehabilitation management extends beyond symptomatic treatment, and emphasises the achievement of the best possible quality of life for the person within the limits of their disease (McGrath and Davis 1992; Schapiro and Langer 1994). This practise is based on clinical judgement rather than scientific evidence since such evidence does not exist.

Aims of inpatient rehabilitation

The primary aims of inpatient neurorehabilitation include:

i) a comprehensive assessment of physical, social and psychological needs

ii) promoting physical, psychological and social adaptation to disability and handicap

iii) facilitating independence in daily activities

iv) maximising life satisfaction for both patient and carers

v) empowerment

iv) preventing secondary complications such as contractures, pressure areas, and pain.
These aims are advocated by leading clinicians in the field of MS rehabilitation (e.g. Langton-Hewer 1980; Slater 1980; Scheinberg et al 1981; Schapiro 1990; La Rocca and Kalb 1992; Thompson 1996c).

**Key elements of the inpatient rehabilitation process**

The key elements of the inpatient rehabilitation process are identified as:

1. **A multidisciplinary team approach**
   
   Inpatient rehabilitation is usually carried out by co-ordinated input from a wide range of professionals including: doctors, continence advisors, nurses, occupational therapists, physiotherapists, psychologists, social workers, and speech and language therapists. This multidisciplinary team approach is believed to be essential in providing an effective service (Erickson et al 1989; ABN Working Party 1992; Greenwood 1992; Barnes 1993; Bakheit 1995).
   
   While the precise composition of this team varies from place to place, a report by the Royal College of Physicians (Physical Disability 1986 and Beyond 1987) identified a number of generic services which were considered to be a central component of any rehabilitation service (Table 1.2)

<table>
<thead>
<tr>
<th>Table 1.2 Services central to a rehabilitation service</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical, occupational and speech therapy</td>
</tr>
<tr>
<td>Continence service</td>
</tr>
<tr>
<td>Orthotic service</td>
</tr>
<tr>
<td>Pressure sore policy and service</td>
</tr>
<tr>
<td>Prescription and maintenance service for basic wheelchairs</td>
</tr>
<tr>
<td>Counselling service</td>
</tr>
<tr>
<td>Effective communication between hospital and community care services</td>
</tr>
</tbody>
</table>
ii) Programmes tailored to meet the needs of the individual

Recognition of the wide ranging and ever-changing needs of people with MS is crucial to their successful health care management (La Rocca and Kalb 1992; Thompson et al 1994). Consequently, common practise is to tailor the goals and process of rehabilitation to meet the needs of the individual at a specific point in time (Vaney 1991; ABN Working Party 1992).

iii) Patient centred functional goal setting approach

Patient centred functional goal setting is believed to be fundamental in ensuring effective rehabilitation (Davis et al 1992; Barnes 1993; Rossiter and Thompson 1995a; Edwards 1996).

Differences in service provision and delivery however do exist (BSRM Working Party Report on MS 1993). The structure and funding mechanism of the health care system appears to have a major influence on the model of care provided in each country. For example, in some countries patients are routinely offered inpatient treatment on an annual basis (Vaney 1991), whereas in others patients are selectively admitted according to current needs (Johnson and Thompson 1996). Differences in the availability of specialist MS facilities (inpatient and outpatient) are also apparent. These were highlighted in a survey undertaken by the International Federation of Multiple Sclerosis Societies (Paty 1994). Of the 26 countries surveyed, 25 countries offered general consultation services, but only 16 provided speciality outpatient care, and 11 specialist inpatient care.
1.4 The need for evaluation of health care services

Despite the seemingly large amounts of resources being devoted to the management of MS, a number of reports have effectively highlighted inadequacies of services and pre-existing unmet need within this group (NHS Community Care Act 1989; Neurological Provision-key areas and targets 1991; ABN Working Party Report 1992; BSMR Working Party Report on MS 1993; Williams and Bowie 1993; Thompson et al 1997, in press). Thompson and colleagues (1997, in press) state that people with MS have been poorly served by the health service to date and potentially have much to gain from the NHS reforms. The thrust behind many of these reforms is to evaluate the effectiveness of health care interventions to ensure people are provided with the most effective and efficient services available (The Health of the Nation 1991).

It is now widely accepted that determining the effectiveness of an intervention by measuring its effect on outcome provides a basis for evidence-based clinical decision making (Hopkins and Costain 1990; Research and Development: towards an evidence-based health service 1995; Sackett et al 1996). Traditionally it has been common to evaluate the outcome of interventions, such as drug therapies in MS, on the basis of entities such as mortality, relapse rate, and lesion load as determined by magnetic resonance imaging. Recently there has been an increasing recognition of the need to address disability, handicap and other aspects of quality of life in both the provision and evaluation of health and social service interventions for people with MS (Rodriguez et al 1994; Devinsky 1995; Johnson 1996; Hobart et al 1996a).
Although anecdotal evidence exists to support the success of inpatient rehabilitation, there is little scientific evidence to substantiate its effectiveness or measure its cost (Tallis 1989; Schapiro 1990; BSRM Working Party Report on MS 1993). The majority of the literature is based on descriptions of the aims and process of rehabilitation and is reliant on subjective evaluations of outcome (e.g. Langton-Hewer 1980; Erickson et al 1989; La Rocca and Kalb 1992). Few studies provide objective data. Such evidence alone is inadequate.

Effectiveness needs to be established in a scientifically acceptable manner (Johnson 1996). There is an urgent need for further outcome studies of comprehensive inpatient rehabilitation to assess and predict outcomes more accurately. The aim of this thesis is to evaluate the effectiveness of inpatient rehabilitation of patients in the progressive phase of MS.
Chapter 2

Measuring the effectiveness of inpatient rehabilitation: what outcomes to measure

In carrying out an evaluation of inpatient rehabilitation in MS, it is necessary to select relevant outcomes to measure, and identify appropriate instruments with which to measure them. This requires a thorough knowledge of the aims of rehabilitation, and what can be expected to change as a result of the intervention. Traditionally, clinical outcomes have been assessed on the basis of simple and easy to measure entities such as mortality, presence or absence of disease, duration of survival, and length of disease-free interval. These assessments are inadequate and largely irrelevant, in the evaluation of health-care interventions for chronic and progressive diseases, such as MS, where the most burdensome consequence of the disease is disablement (Wood 1889), and the primary aim of intervention is to impact on the quality rather than the quantity of life (Hobart et al 1996a). The measurement of the long term consequences of disease has, however, presented both technical and conceptual problems to health care evaluation (Thuriax 1989). The International Classification of Impairments, Disabilities and Handicaps (ICIDH) was developed by the World Health Organisation (WHO) in response to the need for a conceptual and operational framework for describing and measuring the consequences of disease (WHO 1980).
2.1 The ICIDH: a framework for measuring the outcome of inpatient rehabilitation

In the early 1970's the WHO attempted to investigate how different countries were addressing the problems associated with disablement. The collation of information was impossible since neither a common terminology, nor an agreed set of concepts from which it might grow, was in use (Robinson 1985). To solve this problem, they considered that the health-related consequences of disease needed to be classified and standardised in a manner parallel to the International Classification of Diseases which had been developed by the WHO in 1977 (Wood 1989). While recognising that the problems faced by an individual did not fall into neatly defined sections, it was felt that the creation of categories would enable data to be organised in a structured manner which could be understood universally. The aim was to facilitate some coherence in thinking in an area which was arbitrary and disjointed. It was anticipated that this would lead to better understanding and communication, thereby resulting in improved services (de Kleijn-de Vrankrijker et al 1989; Badley 1993). As a result, by 1980 the WHO had developed and published an international classification scheme relating to the consequences of disease, the ICIDH (WHO 1980).

2.1.1 The conceptual framework

The fundamental concept behind the ICIDH is that disease can be defined in relation to its consequences (Ebrahim 1990). This is based on the view that the disease process may be described at different points in its progression:
• from its cause or origin (aetiology)
• to its active state (pathology)
• to the long-term consequences of organic function (impairment)
• in terms of functional change (disability)
• from the perspective of socioeconomic limitations (handicap) (Chamie 1990).

The crux of the ICIDH is that it classifies the consequences of disease into three conceptual categories; impairment, disability and handicap. The conceptual distinctions between the different planes of the consequences of disease is crucial (Badley 1993). The particular relevance of the conceptual distinctions is that they correspond to the obligations of different sectors of the overall care of the individual, with impairments being the primary concern of medical services, disabilities of rehabilitation services, and handicaps of social welfare provisions and broader areas of social policy such as education, employment, transport and housing (Wood 1989). It thereby extends the dimensions of health-related experience previously covered by the concept of disease.

2.1.3 Terminology

Detailed information about each of these dimensions is published in the ICIDH manual (WHO 1980). These are described, in the context of health experience, as follows:

*Impairment* describes "a loss or abnormality of psychological, physiological, or anatomical structure or function". These are disturbances at the level of the
organ. Examples include: blindness, deafness, weakness, spasticity, dysarthria, detrusor instability, memory disturbance and depression.

Disability describes "any restriction or lack (resulting from an impairment) of ability to perform an activity in the manner or within the range considered normal for a human being". These are disturbances in function or activity at the level of the person. Examples include difficulty with: seeing, hearing, speaking, continence, dressing, bathing, shopping, cooking, walking and climbing stairs.

Handicap describes "any disadvantage for a given individual, resulting from an impairment or a disability, that limits or prevents the fulfilment of a role that is normal (depending on age, sex and cultural factors) for that individual". These describe limitations in socio-economic terms. Examples include: social isolation, unemployment, being homebound, and enduring financial difficulties.

The ICIDH uses the term 'disablement' to embrace all three categories, thus referring to all the consequences of disease.

2.1.3 Application to clinical practice

The ICIDH system was designed:

- as a classification scheme for health-related problems
- to provide a theoretical framework for disablement
- in the testing of different models of disablement (e.g. the social, economic or medical models) (Badley 1987).

It was therefore developed in a manner which allows the classifications to be studied either on their own, or in conjunction with any or all of the
classifications, in order to explore the inter-relationship between the different dimensions (Chamie 1990).

The ICIDH has been used increasingly throughout the world in a variety of different fields (including institutions, population surveys and administration issues), and for a variety of purposes (e.g. linguistic agreement and debate, development of the interdisciplinary field of disability, planning and formulation of disability policy, programme evaluation and development, record keeping, survey research and database development) (Badley 1987, 1993; Chamie 1989; Wood 1989; Thuriax 1989; Rodriguez et al 1994). It has been successfully applied to a number of areas specifically within the field of rehabilitation (e.g. fifteen separate studies were cited in de Kleijn-de Vrankrijker et al 1989).

Difficulties have been reported on the use of the ICIDH coding system for evaluation or data collection purposes. These include reports of unreliability, the time consuming nature of assessment (de Kleijn-de Vrankrijker et al 1989), difficulties in grading severity of problems (Badley 1987), and disagreements in the linking of concepts to specific operational definitions (Chamie 1990).

The concepts of the ICIDH are also not without their problems. There is little debate concerning impairment since a consensus appears to exist about what constitutes a loss or abnormality of psychological, physiological or anatomical structure or function. Confusion and disagreement arises however as the focus changes to the more social and cultural aspects of disablement, such as disability and especially handicap. This disagreement occurs in two main areas: in making the distinction between these two concepts; and in the relative value of using disability and handicap as concepts (Badley 1987). Some argue...
that assessments of this kind involve making value judgements as to whether
the disabled person is at a disadvantage in society, with considerable debate
about what is "normal" and the relevance of deviations from the so-called norm
(Badley 1987, 1993; Finkelstein 1989). It is also queried whether it is within the
remit of the health care system to "interfere" with these social aspects of life,
with some refuting this medicalisation of disablement (Finkelstein 1989; French
1992; Johnston 1996). These issues are extremely complex and, although very
important, it is not appropriate to explore them further in the context of this
thesis. It is, however, noteworthy that the ICIDH has been criticised on the
grounds that it institutionalises disablement.

2.1.4 Application of the ICIDH to outcome studies in
rehabilitation

Many outcome studies in rehabilitation lack a theoretical basis, and as a
consequence their formulation and interpretation has been impaired (Wagner
1987). The ICIDH provides a theory of both disablement and the rehabilitation
process. It has proven relevant, amenable to operationalisation, and
reasonably comprehensive from the perspective of rehabilitation aims (Chamie
1990). Each concept has the capacity to be defined, influenced and measured.
It's well established terminology enables the information gathered to be
universally understood. Consequently it is now considered by many to be the
cornerstone of the medical and rehabilitation management of MS (Pearce and
Kirchner 1995). The conceptual basis and terminology of the ICIDH underpins
the theoretical and operational background to the studies undertaken in this thesis.

2.2 Using appropriate measurement instruments

The quality of outcome data is determined by the quality of the instruments used to produce it. Consequently considerable care is required in choosing appropriate measurement instruments. To ensure sound measurement of clinical outcomes it is essential that the instruments chosen have been comprehensively evaluated in terms of both clinical usefulness and scientific properties (Hobart et al 1996b). In accordance with the measurement standards for interdisciplinary medical rehabilitation (Johnston et al 1992) these factors were a prime consideration in choosing which instruments to use in the studies undertaken in this thesis.

For the purposes of these studies, clinical utility required that the instruments adhered to the concepts described by the ICIDH, were user friendly, practical to administer, inexpensive, and could be undertaken within both an inpatient and outpatient setting. In terms of scientific soundness, three essential properties were required: reliability, validity, and responsiveness. These psychometric properties are key features in the evaluation of all measurement instruments. Knowledge about them is essential to enable accurate interpretation of the information gained (Streiner and Norman 1989; Wilkin et al 1992). These features, which will be referred to frequently throughout the remainder of this chapter, are described briefly below.
Reliability

A reliable instrument produces results which are accurate, consistent, stable over time, and reproducible. There are three types of reliability relevant to the evaluation of measures used in this thesis:

i) Internal consistency - the extent to which items comprising a scale measure the same concept.

ii) Test-retest reliability - the stability of an instrument over time. This determines whether the instrument used on the same subject on two different occasions provides the same results.

iii) Rater reliability - the agreement within a single rater (intrarater) or between two or more raters (interrater). This type of reliability is not relevant for self-administered instruments.

Each different type of reliability determines different sources of error, and therefore each must be examined to ensure a comprehensive evaluation (Streiner and Norman 1989).

Validity

A valid instrument is one which measures what it is intended to measure (Wilkin et al 1992). The term refers to the appropriateness, meaningfulness, and usefulness of a measure and of the inferences made from it (Johnston et al 1992). There are three basic interrelated types of validity:

i) Content validity - the extent to which a measure is representative of the conceptual domain it is intended to cover.

ii) Criterion-related validity - the degree to which a measure correlates with a gold standard (the criterion).
iii) **Construct validity** - the degree to which an instrument measures the theoretical construct it was designed to measure. Here evidence is gathered by series of studies to support the validity of the instrument. A series of questions are asked to determine whether the instrument measures what it is supposed to measure (convergent construct validity); does not measure what it does not intend to measure (discriminant construct validity); distinguishes between groups in predictable ways (group differences construct validity); and produces results consistent with theoretical expectation (hypothesis testing).

Each type of validity provides a different approach to assessing the extent to which an instrument measures what it purports. All three types of validity should be considered for a comprehensive evaluation (Hobart et al 1996b).

**Responsiveness**

Responsiveness is the ability of the instrument to detect minimal clinically significant change (Guyatt et al 1987). This feature is especially important when evaluating the effectiveness of health care interventions (Fitzpatrick et al 1992a). Unresponsive measures may not be capable of detecting the changes which have occurred as a result of the intervention.

The validity of outcomes research is dependent upon both the scientific rigour of the trial design and the validity of its measures (Johnston and Granger 1994; Rudick et al 1996). Knowledge of the psychometric properties of the instruments used is therefore essential. Ideally the instruments should have been comprehensively evaluated, and possess, all of these scientific attributes.
Unfortunately however, very few of the instruments available for measuring the clinical outcome of rehabilitation, particularly in progressive diseases such as MS, achieve this ideal (Ebrahim 1990). In reality compromises often have to be made when choosing an instrument. By understanding the strengths and weaknesses of instruments, appropriate cautions and reservations can be made when interpreting study results (Johnston et al 1992). As a consequence considerable emphasis has been placed on the clinical and scientific aspects of the measures used throughout this thesis.

2.3 Measuring Impairment

Neurological impairments are the direct physiological consequences of the underlying pathology. They are the clinical signs and symptoms produced by damage to the nervous system. In MS, where any part of the CNS may be affected, they are many and varied. The Association of British Neurologists Working Party on MS (1992) described over 20 different impairments which may be experienced (Table 2.1).

Table 2.1 Range of impairments experienced by people with MS

<table>
<thead>
<tr>
<th>Poor co-ordination / ataxia</th>
<th>Weakness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spasticity</td>
<td>Fatigue</td>
</tr>
<tr>
<td>Loss of proprioception</td>
<td>Visual disturbances</td>
</tr>
<tr>
<td>Communication difficulties</td>
<td>Swallowing difficulties</td>
</tr>
<tr>
<td>Urinary disturbances</td>
<td>Bowel disturbances</td>
</tr>
<tr>
<td>Sensory disturbances</td>
<td>Sexual dysfunction</td>
</tr>
<tr>
<td>Pain</td>
<td>Dizziness and unsteadiness</td>
</tr>
<tr>
<td>Cognitive dysfunction</td>
<td>Depression and anxiety</td>
</tr>
<tr>
<td>Oedema</td>
<td>Paroxysmal symptoms</td>
</tr>
</tbody>
</table>
In the past the level of impairments was used as an index to indirectly reflect the severity and extent of the disease process. Since the aim of medical intervention was to halt or cure this process, impairment was the focus for the evaluation of health care interventions. Studies now demonstrate that neither measures of impairment nor disability, as determined by clinical neurological examination, correlate strongly with the Magnetic Resonance Imaging lesion load (thought to be a more direct although non specific index of pathology) either in the cerebrum or spinal cord (Thompson et al 1990; Miller 1994). Although more pathologically specific measures such as spinal cord atrophy show a stronger correlation and therefore look promising (Losseff et al 1996). Consequently it is now considered that impairment instruments rate the severity of the presenting signs and symptoms as determined by the standard neurological examination, rather than the pathologic extent of the disease itself (Willoughby and Paty 1988; Wade 1996).

Nevertheless it remains important to rate impairment for a number of reasons:
- it is currently the general standard for quantifying the severity of the disease process, albeit established indirectly, since no generally accepted laboratory measure of disease activity or severity for MS exists (Willoughby and Paty 1988; Sharrack and Hughes 1996)
- it is the most relevant outcome to measure in interventions which aim to directly affect specific impairments (e.g. drug intervention to reduce spasticity)
- by providing a reflection of the current level of disease severity, it helps to establish whether the outcome is attributable to the intervention or to
spontaneous neurological improvement (or deterioration). This is particularly relevant in the evaluation of interventions for MS where the disease process is variable and unpredictable.

2.3.1 Instruments for measuring impairment in MS

Over the past 35 years a number of instruments have been proposed for rating the severity of neurological impairment in MS (Table 2.2).

<table>
<thead>
<tr>
<th>Author</th>
<th>Title</th>
<th>Introduced</th>
</tr>
</thead>
<tbody>
<tr>
<td>*Arkin et al</td>
<td>Arkin Scale</td>
<td>1950</td>
</tr>
<tr>
<td>*Alexander</td>
<td>-</td>
<td>1951</td>
</tr>
<tr>
<td>*Kurtzke</td>
<td>Disability Status Scale</td>
<td>1955</td>
</tr>
<tr>
<td>Kurtzke</td>
<td>Functional Systems</td>
<td>1961</td>
</tr>
<tr>
<td>Kurtzke</td>
<td>Expanded Disability Status Scale</td>
<td>1983</td>
</tr>
<tr>
<td>*Fog</td>
<td>-</td>
<td>1965</td>
</tr>
<tr>
<td>Potvin et al</td>
<td>Quantitative Neurological Examination</td>
<td>1975</td>
</tr>
<tr>
<td>*Patzold et al</td>
<td>-</td>
<td>1978</td>
</tr>
<tr>
<td>*Mickey et al</td>
<td>Illness Severity Scale</td>
<td>1984</td>
</tr>
<tr>
<td>*Sipe et al</td>
<td>Neurologic Rating Scale (Scripps Scale)</td>
<td>1984</td>
</tr>
<tr>
<td>Confavreaux et al</td>
<td>EDMUS Impairment Scale</td>
<td>1992</td>
</tr>
<tr>
<td>Hohol et al</td>
<td>Disease Steps Scale</td>
<td>1995</td>
</tr>
<tr>
<td>Starr et al</td>
<td>MS Symptom Inventory</td>
<td>1996</td>
</tr>
<tr>
<td>Ravnborg et al</td>
<td>MS Impairment Scale</td>
<td>1996</td>
</tr>
</tbody>
</table>

* cited in Willoughby and Paty 1988

Of these instruments, Kurtzke's Expanded Disability Status Scale and Functional Systems, designed specifically for the MS population, is distinguished from the others by virtue of its widespread use (Rudick et al 1996). It is generally regarded as the preferred measure of treatment outcome in clinical trials (Goodkin 1996), is the most widely used method of clinical
assessment in MS (Whitaker et al 1995; Sharrack and Hughes 1996; Weinshenker et al 1996a), and is commonly used to classify the severity of MS (Quick and Schapiro 1996). It constitutes part of the Minimal Record of Disability (MRD), a system developed by The International Federation of Multiple Sclerosis Societies (IFMSS) in an attempt to standardise information collected on MS (Slater 1984).

2.3.2 Expanded Disability Status Scale and Functional Systems

The Disability Status Scale (DSS), a simple 10 point scale, was originally devised by Kurtzke in 1955 to evaluate isoniazid as a possible treatment in MS. This instrument was extended in 1961 to incorporate the assessment of different functional groups within the CNS termed Functional Systems (FS) (Kurtzke 1961), thereby forming a two part rating instrument. The intention of these instruments, each complementary to the other, was to measure neurological change over time on the continuum of normal to maximal disease. Further refinement of the FS instruments occurred in 1965 (Kurtzke 1965) and then again in 1983 at which time the Disability Status Scale was also developed and renamed as The Expanded Disability Status Scale (EDSS, Kurtzke 1983).

The Functional Systems

The FS were specifically designed to delineate the nature and severity of neurologic insults caused by MS. Comprised of eight complementary instruments, each assesses a specific functional system within the CNS and is
defined in terms of neuroanatomical area: Pyramidal, Cerebellar, Brainstem, Sensory, Bowel and bladder, Visual, Cerebral (or mental), and Other (miscellaneous systems). Together these systems comprise all neurological abnormalities on examination which can be attributed to MS (Kurtzke 1983). Kurtzke (1983) states that each instrument is mutually exclusive in terms of neuroanatomy. This is disputed by Willoughby and Paty (1988) who argue that "a single lesion in the brainstem may produce dysarthria (brainstem function) together with a weakness in limbs (pyramidal function) and a single lesion in the spinal cord may produce pyramidal tract dysfunction plus impairment of sensation".

Only objectively verifiable signs due to MS are said to be included in the instruments with symptoms being discarded (Kurtzke 1983). This is disputed by Willoughby and Paty (1988) who state that the grading in some functional systems is determined by symptoms described by the patient (e.g. bowels and bladder) and therefore are not verified objectively.

Seven of the eight systems are graded from zero (normal) to five or six (maximal impairment). The one exception is the 'Other' category which is dichotomous, with zero as none and one as present. The higher the score, the greater is the dysfunction (Appendix 1).

The FS are not additive, each system can only be compared over time with itself. Kurtzke (1983) provides a clinical example as to why this should be the case; "as pyramidal worsens, cerebellar will ‘improve’, since patients cannot be
ataxic if they cannot move”. This lack of additivity of the FS was the reason for
the introduction of the EDSS which, for evaluation purposes, can be used to
provide an overall comparison of the same patient at different points in time, or
comparison within or among groups of MS patients over a period of time.

The Expanded Disability Status Scale

The EDSS is based on the standard neurological examination and is intended
to rate the maximal function of each patient as limited by his/her neurological
deficits (Kurtzke 1983, refer to Appendix 2). It rates the overall neurological
status across a 20 point scale from zero (normal) to 10 (death due to MS).

The EDSS provides an index of neurological deficit. By weighting individual
scores for each Functional System in combination with additional data, based
largely on gait dysfunction, it forms an overall score. From grades zero to three
the steps are defined by variations in impairment as determined by the FS
scores. From grades four to seven, the steps are largely dependent on one
aspect of disability, locomotion. From grades seven to ten the focus is on upper
limb and brainstem function. While its basic premise is that “it permits an
equivalent grade for what is hoped to be an equivalent amount of disease
regardless of the structures involved” (Kurtzke 1961), its ability to achieve this
has been questioned (Willoughby and Paty 1988; Ellison et al 1994). For
example, it has been criticised for its heavy bias in the assessment of gait, and
its lack of assessment of cognitive and upper limb disabilities in a disease
which randomly affects any part of the CNS (Willoughby and Paty 1988; Cohen
Kurtzke (1965,1983) has suggested that the results of the EDSS follow a Gaussian appearance (normal distribution). However several cross sectional studies have shown an unequal bimodal distribution, with peaks between 1.0 to 3.5 and 6.0 to 6.5, and a relative scarcity of patients at grades 4.0 to 5.5 (Dassel 1981; Goodkin et al 1992; Gulick et al 1993; Whitaker et al 1995; Hohol et al 1995). Kurtzke (1983) suggests that the likely cause for this is that the examiners had not assigned the scores in the lower range of the scale with regard to the FS, according to the scoring procedure.

More recently the EDSS scale has shown to be non-linear in nature (Ellison et al 1994; Weinshenker et al 1996a). Weinshenker and colleagues (1996a) demonstrated that patients who scored between 3.0 to 5.0 on the scale were much more likely to worsen by one EDSS point than those who scored between 6.0 to 7.0. Their predictive models of short term progression indicated that the probability of worsening was best measured in terms of the distance from a score of 4.5, the disability level at which chances of progression were greatest. The probability of worsening was less at both higher and lower EDSS levels. Furthermore they showed that the duration of MS was significantly associated with observed worsening by one EDSS point. Patients with a disease duration of 20 years or less were more likely to deteriorate than those with a longer disease duration.
2.3.3 Psychometric properties of the EDSS and FS

Despite the extensive use of the EDSS and FS in a variety of clinical trials (e.g. O’Khatri et al 1985; Paty et al 1993; Johnson et al 1995; Jacobs et al 1996), surprisingly few studies have investigated the psychometric properties of these instruments in any detail. Those undertaken fail to provide clear and comprehensive scientific evidence of the reliability, validity or responsiveness of these instruments. In addition, methodological differences between the studies, such as variations in sample size and rating procedures, make it difficult to compare and contrast their results, which are wide-ranging and sometimes contradictory. Conclusions should remain tentative until further work is undertaken.

1) Reliability

Few multi-centre trials of treatment in MS have considered intercentre, interobserver, and intraobserver discrepancies as potential sources of error in comparisons of groups of patients receiving different treatments (Francis et al 1991). Evaluation of its reliability, in accordance with accepted guidelines (Medical Outcomes Trust 1995), remains incomplete. For example, no evidence regarding test-retest reliability is reported.

Interrater reliability

Evidence concerning the interrater reliability of Kurtzke’s Scales is variable and sometimes conflicting. The majority of studies have used the Kappa coefficient (K) to determine the extent of agreement among raters to be expected by chance. Conclusions in terms of exact agreement vary from poor to almost perfect for the EDSS (Table 2.3), and slight to almost perfect for the seven
components of the FS (Table 2.4). There is a general agreement that reliability is lower when the EDSS score is less than five (Amato et al 1988; Verdier-Taillefer et al 1991; Hohol et al 1995), although Noseworthy and colleagues (1990) contradict this suggesting that no single step on these instruments is particularly prone to variability.

Moderate discrepancies in agreement might be expected, given the complexity and variability of the MS population, and where clinical features such as fatigue may interfere with consecutive neurological examinations. Differences in methodology may also contribute to some of this variation. For example, whereas an initial study by Amato in 1987 reported the interrater reliability to be high, on conducting a second study in 1988, it was found to be an unacceptably low 30-50%. The authors suggested that differences in the examination procedures of each study accounted for this variation. In the first study each patient was examined by one neurologist in the presence of three others, following which all four rated the patient on the EDSS and FS. In contrast, in the second study four neurologists, arranged into six pairs, each examined and rated the patients on the scales, thereby taking into account the variability inherent in the clinical examination. It was suggested that the lower level of reliability, evidenced in the second study, was likely to be more indicative of reliability within the normal clinical setting. A study by Francis et al (1991) further substantiated this finding.
<table>
<thead>
<tr>
<th>Study</th>
<th>Year</th>
<th>N</th>
<th>EDSS range</th>
<th>Method</th>
<th>Level of exact interrater agreement: Kappa co-efficient (K).</th>
<th>Level of agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amato et al</td>
<td>1987</td>
<td>40</td>
<td>1.0 - 9.0</td>
<td>1 neurologist examines the patient in the presence of 3 others, and each then score the patient independently</td>
<td>0.95</td>
<td>very high</td>
</tr>
<tr>
<td>Amato et al</td>
<td>1988</td>
<td>24</td>
<td>1.0 - 8.5</td>
<td>4 neurologists arranged into 6 pairs, each assessing the patient</td>
<td>0.49</td>
<td>moderate</td>
</tr>
<tr>
<td>Noseworthy et al</td>
<td>1990</td>
<td>168</td>
<td>4.0 - 6.5</td>
<td>2 physicians rating 168 patients in 545 paired examinations (blinded to previous assessments and each others scores)</td>
<td>0.62</td>
<td>substantial</td>
</tr>
<tr>
<td>Verdier-Tailefer et al</td>
<td>1991</td>
<td>59</td>
<td>0.0 - 8.5</td>
<td>2 centre trial: patients assessed independently by 2 neurologists, and pairs of neurologists</td>
<td>EDSS &lt; 5: 0.15 - 0.18 EDSS &gt;= 5: 0.38 - 0.44</td>
<td>poor</td>
</tr>
<tr>
<td>Francis et al</td>
<td>1991</td>
<td>20</td>
<td>3.0 - 9.5</td>
<td>Patients assessed independently by 3 separate neurologists</td>
<td>0.32 - 0.76</td>
<td>fair - substantial</td>
</tr>
<tr>
<td>Goodkin et al</td>
<td>1992</td>
<td>10</td>
<td>1.0 - 3.5</td>
<td>Patients assessed by 4 trained examining physicians</td>
<td>ICC 0.654 - 0.708 N.B. Intra Class Correlation Co-efficient (ICC) used not Kappa</td>
<td></td>
</tr>
<tr>
<td>Hohol et al</td>
<td>1995</td>
<td>60</td>
<td>1.0 - 9.0</td>
<td>Patients examined independently by 2 neurologists (within one hour)</td>
<td>0.54</td>
<td>Moderate</td>
</tr>
<tr>
<td>Hobart et al</td>
<td>1996c</td>
<td>116</td>
<td>5.0 - 9.0</td>
<td>Patients examined independently by 2 neurologists</td>
<td>ICC = 0.77 N.B. ICC used not Kappa</td>
<td></td>
</tr>
</tbody>
</table>

The reliability level as a function of the K value is judged by the reference values suggested by Landis and Koch (1977): <0= poor, 0.00-0.20= slight, 0.21-0.40= fair, 0.41-0.60= moderate, 0.61-0.80= substantial, 0.81-1.00= almost perfect.
Table 2.4 Studies evaluating FS interrater reliability

<table>
<thead>
<tr>
<th>Study</th>
<th>Year</th>
<th>N</th>
<th>EDSS range</th>
<th>Method</th>
<th>Level of exact agreement - K Coefficient according to each Functional System</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyramidal</td>
<td>Cerebellar</td>
</tr>
<tr>
<td>Amato et al</td>
<td>1987</td>
<td>40</td>
<td>1.0-9.0</td>
<td>1 neurologist examined the patient in the presence of 3 others, each then scored the patient independently</td>
<td>0.90</td>
</tr>
<tr>
<td>Amato et al</td>
<td>1988</td>
<td>24</td>
<td>1.0-8.5</td>
<td>4 neurologists arranged into 6 pairs</td>
<td>0.28</td>
</tr>
<tr>
<td>Noseworthy et al</td>
<td>1990</td>
<td>168</td>
<td>4.0-6.5</td>
<td>2 physicians rating 168 patients in 545 paired examinations (blinded to previous assessments and each others scores)</td>
<td>0.47</td>
</tr>
<tr>
<td>Verdier-Tailefer et al</td>
<td>1991</td>
<td>59</td>
<td>0.0-8.5</td>
<td>2 centre trial: patients assessed independently by 2 neurologists, and pairs of neurologists</td>
<td>0.45</td>
</tr>
<tr>
<td>Francis et al</td>
<td>1991</td>
<td>20</td>
<td>3.0-9.5</td>
<td>Patients assessed independently by 3 separate neurologists</td>
<td>0.28 - 0.36</td>
</tr>
<tr>
<td>Goodkin et al</td>
<td>1992</td>
<td>10</td>
<td>1.0-3.5</td>
<td>Patients assessed by 4 trained examining physicians</td>
<td>ICC 0.423 - 0.645</td>
</tr>
</tbody>
</table>

The reliability level as a function of the K value is judged by the reference values suggested by Landis and Koch (1977): <0 = poor, 0.00-0.20 = slight, 0.21-0.40 = fair, 0.41-0.60 = moderate, 0.61-0.80 = substantial, 0.81-1.00 = almost perfect.
It is important to consider what level of agreement is acceptable as a reliable indicator of clinical change. Differences are evident in the reliability studies undertaken. For example, in the EDSS "almost perfect" agreement was reported when the criterion was one point \((K=0.94, \text{Amato et al 1988; } K=0.89, \text{Noseworthy et al 1990})\). This level of agreement was "perfect" when the criterion increased to a difference of two points \((K=1.0, \text{Amato et al 1988})\). Similarly in the FS, reported \(K\) values ranged from 0.58-1.0 (moderate to substantial agreement) when the criterion of agreement was one point, with even higher \(K\) values of 0.87-1.0 (almost perfect to perfect) when the criterion was extended to two points (Amato 1988). In contrast, Goodkin et al (1992) found that when a number of different examiners rated the patient, reproducible scoring could only be accomplished within the larger range of 1.5 EDSS or three individual FS points.

**Intrarater reliability**

Intrarater reliability has been demonstrated as very high within one EDSS point (Amato 1987; Goodkin et al 1992; Hobart et al 1996c), and two individual FS points (Amato 1987; Goodkin et al 1992).

In summary, the diversity of results across centres casts doubt on the reliability of these instruments (Willoughby and Paty 1988; Noseworthy et al 1989; Whitaker et al 1995). It has therefore been recommended that outcome studies should:

- include reliability data for all examining physicians,
- only report a change in the degree of disability or response to treatment if a
change of at least one point on the EDSS and two points on the FS has occurred (Amato et al 1988; Noseworthy et al 1989, 1990; Verdier-Taillefer et al 1991; Whitaker et al 1995),

- whenever possible use a single examiner to assess all patients (Goodkin et al 1992).

2) Validity

Evidence to support the validity of Kurtzke’s instruments is scant. Only one aspect of validity, convergent construct validity, has been formally investigated. The evidence available for the EDSS is somewhat confusing (Willoughby and Paty 1988). Some studies validate it as a measure of impairment (Gulick et al 1993), while others validate it as a measure of disability (Cohen et al 1993). Examination of the EDSS reveals a mixture of impairment and disability components (Willoughby and Paty 1988; Sharrack and Hughes 1996), with Kurtzke describing it as measuring each of these dimensions (Kurtzke 1961, 1965, 1970, 1983). No such confusion appears to exist regarding the FS scales which are widely accepted as measuring impairment (Slater 1984; Gulick et al 1993)

Convergent construct validity

Convergent construct validity determines the extent to which the EDSS and FS correlate with other measures of the same construct. Evidence to support the EDSS as a measure of disability includes strong correlation with other disability
instruments including the: Activities of Daily Living Scale \( (r = 0.82, p<0.001, \) Cohen et al 1993\); total score on the Functional Independence Measure \( (r = -0.91, p < 0.0001, \) Brosseau and Wolfson 1994\); Incapacity Status Scale \( (r >0.8, \) Fog et al 1984\); Barthel Index \( (r = 0.89, \) Hobart et al 1996c\); Guy’s Neurological Disability Scale \( (r = 0.77, p <0.001, \) Sharrack et al 1996\). It has also been shown to correlate highly with patient self reported activities of daily living \( (72\% \text{ agreement}, \) Gulick et al 1993\), and to a moderate level with the locomotor item on the Functional Independence Measure \( (r = -0.58, p < 0.0001, \) Brosseau and Wolfson 1994\) and the patient report visual analogue scale of disability \( (Hobart et al 1996c)\).

With regard to its validity as a measure of impairment, the DSS and FS have shown to reflect closely the standard clinical neurological examination \( (Kurtzke 1965; Rose et al 1970)\) and patient self reported symptoms \( (73-86\% \text{ agreement among the FS scales, Gulick et al 1993})\).

**Discriminant construct validity**

High correlations are not expected between scales which are not designed to measure the same construct. The EDSS correlated less well with six non-disability measures with correlations ranging from \( r = 0.02 \) \( (\) Short-form 36 change in health dimension) to \( r = 0.63 \) \( (\) Waterlow pressure sore risk score) \( (Hobart et al 1996c)\).
3) Responsiveness

No studies formally investigate the responsiveness of the FS. One study has investigated responsiveness of the DSS (Ellison et al 1993) and one the EDSS (Hobart et al 1996c). Both of these are published in abstract form. Ellison and colleagues (1993) compared the neurologist's clinical assessment of neurological status (the “gold standard”) with changes in DSS scores, finding the responsiveness to be poor. The EDSS was also reported to be insensitive in detecting change in a population of rehabilitation inpatients (Hobart et al 1996c).

In support of the use of the EDSS and FS as indicators of change, Kurtzke (1961, 1983) demonstrated how scores varied between hospital admissions and discharge, claiming it to be ‘reasonably’ sensitive to changes in treatment trials in acute bouts of the disease process (Kurtzke 1983). Others however dispute this, criticising the EDSS for being insensitive to changes in studies of chronic MS, particularly in the middle ranges of the scale between 6.0-7.5 (Willoughby and Paty 1988; Noseworthy et al 1990; Weinshenker et al 1996a).

In summary, Kurtzke’s EDSS and FS scales were chosen for measuring the level of neurological severity in the studies presented in this thesis. Although reservations have been expressed regarding the psychometric properties of these instruments, currently no well evaluated alternative is available. To date, the EDSS has been used as a key outcome measure in the vast majority of clinical trials in MS.
2.4 Measuring Disability

Just as the severity and diversity of impairments in people with MS is wide ranging, so too is the spectrum of disabilities, varying from total dependence to independence in all self care activities (BSRM Working Party Report on MS 1993). Table 2.5 describes just a few of the many disabilities experienced in MS. Warren (1989) estimated that approximately 83% of people with MS experience some level of disability.

Table 2.5 Range of disabilities and their estimated prevalence in the UK

<table>
<thead>
<tr>
<th>Disability</th>
<th>Estimated prevalence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Locomotion difficulties</td>
<td>* 45% unable to walk unaided</td>
</tr>
<tr>
<td></td>
<td>* 30% essential wheelchair users</td>
</tr>
<tr>
<td></td>
<td>* 8% occasional wheelchair users</td>
</tr>
<tr>
<td>Bladder incontinence</td>
<td>≈33% experience dysfunction of bladder at some stage</td>
</tr>
<tr>
<td></td>
<td>** 78% experience a disturbance of micturition at some stage</td>
</tr>
<tr>
<td></td>
<td>** 46% develop some degree of urinary incontinence</td>
</tr>
<tr>
<td>Faecal incontinence</td>
<td>≈ 33% experience dysfunction of bowel at some stage</td>
</tr>
<tr>
<td></td>
<td>≈ 10% experience faecal incontinence as a persistent problem</td>
</tr>
<tr>
<td></td>
<td>≈ 4% experience faecal incontinence &gt; once/ fortnight</td>
</tr>
<tr>
<td>Visual disturbance</td>
<td>≈ 60-70% experience disturbances of vision</td>
</tr>
<tr>
<td></td>
<td>≈ 10% have seriously impaired vision and are eligible for registration as partially sighted or blind</td>
</tr>
<tr>
<td>Sexual dysfunction</td>
<td>* 72% of males and 61% of females reported disturbances in sexual activity</td>
</tr>
<tr>
<td></td>
<td>* 35% of males with mild disability (EDSS = 0-4.5) reported impotence</td>
</tr>
<tr>
<td></td>
<td>* 60% of males with moderate disability (EDSS 5.0 - 6.5) reported impotence</td>
</tr>
<tr>
<td>Cognitive deficits</td>
<td>≈20% experience some degree of intellectual dysfunction</td>
</tr>
<tr>
<td></td>
<td>≈30% show signs of impaired memory on detailed testing. This is severe and likely to contribute significantly to disability in 10%</td>
</tr>
<tr>
<td></td>
<td>≈15% show difficulty in the ability to make judgements, and in demonstrating ‘socially acceptable’ behaviour</td>
</tr>
</tbody>
</table>

Traditionally maximising independence in functional activities has been considered a major goal in MS rehabilitation (Fuhrer 1987; Wagner 1987; Granger et al 1990). Consequently disability has been a primary focus for its evaluation. Many clinicians, researchers, and policy makers now believe that systematic functional assessment, through the use of standardised measurement instruments, is an essential component of rehabilitation programmes (e.g. Keith et al 1987; Hall et al 1993; Thompson 1996c). This information is useful in: (a) establishing the level of disability (Keith et al 1987); (b) determining the priorities of the programme (Disler et al 1993); (c) quantifying any changes which may occur following intervention (Hall et al 1993); (d) predicting outcome (Thompson 1996c); (e) planning placement and (f) estimating care requirements (Heinemann et al 1993).

2.4.1 Studies in Multiple Sclerosis

Although the prevalence of specific disabilities is reasonably widely reported (e.g. Roberts et al 1991; Mathews 1991; BSMR Working Party Report on MS 1993) few authors have utilised standardised measurement instruments to investigate disability in MS or have quantified how disability can be influenced by health care intervention. Only eight studies systematically evaluate this outcome in inpatient rehabilitation. These studies are described in detail in chapter three.

Kuroiwa (1984) attempted to determine general trends of disabilities in a Japanese population (n = 75). Using the Incapacity Status Scale (ISS) he found most problems occurred in stair climbing, vision, ambulation and bathing. A
A strong positive correlation was demonstrated with the level of disability and: the duration of illness; impairment of pyramidal ($r = 0.73 - 0.84$) and sphincter function ($r = 0.72 - 0.95$) as measured by Kurtzke's FS; and handicap as measured by the Environmental Status Scale ($r = 0.78 - 0.89$).

A similar pattern of disability was found in a survey undertaken by Johnson and colleagues (1984) in New Zealand. Using the ISS to measure disability, they reported most problems occurring in the areas of stair-climbing, ambulation, bathing and fatiguability.

Granger et al (1990) investigated disability in 24 people with MS representative of a range of severity of disabilities. He attempted to predict the levels of burden of care, and subjective satisfaction with life using combinations of functional assessment scales. In terms of burden of care, people required an average of 63.54 minutes of help each day (sd 16.87; range 0-266) to undertake daily activities. With regard to life satisfaction, five people reported they were very satisfied with life, eight were fairly well satisfied, five were more satisfied than not and six were not satisfied. He demonstrated that the motor domain items of the Functional Independence Measure (FIM) together with a measure of vision were strong predictors of the amount of help required. Prediction of life satisfaction however was dependent on items from a number of the scales (including three FIM items and two handicap items) in association with scores on the Brief Symptom Inventory Scale.
2.4.2 Instruments for measuring disability

Disability scales are commonly used in MS centres and clinics (Ketelaer 1995). Over the past 25 years a variety of instruments has been employed (Table 2.6).

Table 2.6 Instruments used to evaluate disability in MS

<table>
<thead>
<tr>
<th>Measurement Instrument</th>
<th>Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barthel Index</td>
<td>Granger et al 1990; Kidd et al 1995b</td>
</tr>
<tr>
<td>Incapacity Status Scale</td>
<td>Reding et al 1987; Francabandera et al 1988</td>
</tr>
<tr>
<td>Northwick Park ADL Index</td>
<td>Mertin et al 1984</td>
</tr>
<tr>
<td>Burke Stroke Time-oriented Profile</td>
<td>Feigenson et al 1981</td>
</tr>
<tr>
<td>Revised Level of Rehabilitation Scale</td>
<td>Carey et al 1988</td>
</tr>
<tr>
<td>Functional Independence Measure</td>
<td>Brosseau and Wolfson 1994; Granger et al 1990; Kidd et al 1995b</td>
</tr>
<tr>
<td>Guy’s Neurological Disability Scale</td>
<td>Sharrack et al 1996</td>
</tr>
</tbody>
</table>

These scales measure a wide range of activities and behaviours, including aspects of physical function, cognition, communication, and behaviour.

Unfortunately, to date there is a paucity of evidence to support the psychometric properties of many of these scales (McDowell and Newell 1987; Hobart et al 1996b), with little agreement as to the best instrument for evaluating the impact of rehabilitation (Wade 1992).

Although the Incapacity Status Scale (ISS) was developed as part of the Minimal Record of Disability for MS (Slater 1984) it did not meet the criteria for use as an outcome measure in the studies undertaken in this thesis (refer to Section 2.2). Only one study describes its use in evaluating the outcome of comprehensive rehabilitation (Reding et al 1987) and evidence of its psychometric properties is sparse. It has been criticised for a number of issues including its lack of responsiveness (Sharrack and Hughes 1996), ambiguity
and lack of precise definitions (Colville 1984), for mixing impairment, disability and handicap (Mertin et al 1984) and a lack of refinement of its scoring criteria (La Rocca et al 1984). In preference, the Functional Independence Measure was chosen as the measurement instrument.

2.4.3 The Functional Independence Measure

In 1983 a National Task force was established in North America with the role of developing a minimum data set for general rehabilitation. The purpose of the task force was to develop a dataset, which would enable standardised information to be collected at large numbers of hospitals, in order to facilitate larger scale evaluation and research. A measure of disability was considered an essential component of this data set. The task force reviewed the disability literature and concluded that while a number of disability scales were currently in use (e.g. Barthel Index, Incapacity Status Scale), they were inadequate in their assessment of communication and cognitive deficits. Consequently the FIM was developed as a more comprehensive disability measurement instrument (Keith et al 1987).

The FIM is a generic measure of disability. It addresses a range of basic daily functions and is suitable for a wide variety of diagnostic groups. The FIM is designed to be used by any trained clinician regardless of discipline (Dodds et al 1993). Its theoretical premise is that disability can be evaluated in terms of burden of care, as measured by the consumption of social and economic resources (Granger et al 1993).
The FIM is an ordinal scale, comprising 18 items which can be categorised into two basic domains, motor and cognitive functions (Linacre et al. 1994). These domains are further sub-divided into six subscales: self care, sphincter control, transfers, locomotion, communication, and social cognition (Appendix 3). Items are assessed, by observation of performance, on a 7 level scoring system, based on the type and amount of assistance required for a person to perform the activities (Appendix 3). The score reflects the performance rather than the capability of the person to undertake these activities.

Over the past ten years, the FIM has gained widespread acceptance, nationally and internationally as a measure of disability. By 1993, data on over 150,000 patients had been entered into the American Uniform National Data System (Dodds et al. 1993). A growing body of publications referencing the FIM has subsequently accrued, with 87 publications cited in Medline, 65 in Cinahl and 18 in Rehabilitation Index between 1987 and 1996. Such widespread use of a single, standardised scale offers the advantage of comparing results between studies and conditions. Of these studies four have reported its use specifically in the MS population (Granger et al. 1990; Brousseau and Wolfson 1994; Aisen et al. 1996; Kidd and Thompson 1997, in press), and others have included MS patients within their study population (e.g. Disler et al. 1993; Dodds et al. 1993; Heinemann et al. 1994; Kidd et al. 1995a; Freeman et al. 1996a).

2.4.4 Psychometric properties of the FIM

The FIM has undergone reasonably extensive psychometric testing which is detailed in the following section.
1) Reliability

**Internal consistency**

Internal consistency is assessed by a statistical technique, Cronbach's alpha ($\alpha$), with values greater than 0.7 suggesting that items are measuring the same construct (Dodds et al 1993). The FIM has demonstrated high overall internal consistency (Dodds et al 1993, $\alpha$ 0.93; Brosseau and Wolfson 1994, $\alpha$ 0.94; Hobart et al 1996d, $\alpha$ 0.94) suggesting that it measures a unidimensional construct.

**Interrater reliability**

A number of studies have investigated the interrater reliability of the FIM total and sub-scale scores (Table 2.7). The levels of interrater reliability for the motor domain consistently range from good to very good. In contrast, the reliability levels of the FIM cognitive domain are variable. While some studies report good reliability (Hamilton et al 1987; Chau et al 1994), others demonstrate only fair (Segal et al 1993) or poor levels in this domain (Davidoff et al 1990; Grey and Kennedy 1993; Brosseau and Wolfson 1994). The methodological differences in these studies make comparison of their results difficult. Generalisability of findings is limited by their study designs. For example, Grey and Kennedy (1993) specifically studied spinal cord injured patients, a group who do not experience severe communication and cognitive problems. This makes it difficult to determine whether their results are due to ceiling effects or levels of agreement between raters. Further investigation is necessary to clarify the variability of results in the cognitive domain.
| Study                  | Year | Population (n)                                                                 | Method                                                                                                                                                                                                 | Level of interrater agreement (total score)                                                                 | Level of inter-rater agreement (subscales scores)                                                                 |
|-----------------------|------|-------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------|
| Hamilton et al        | 1987 | 303 rehabilitation inpatients: variety of disease conditions                  | Observer pairs rating the same patient on admission (184 observer pairs on discharge).                                                                                                                  | ICC = 0.86 (admission score) ICC = 0.88 (discharge score)                                                                                                     |
| Hamilton et al (Abstract) | 1991 | 263 rehabilitation inpatients: variety of disease conditions                  | Two or more pairs of clinicians rating 263 patients at 21 centres.                                                                                                                                 | ICC = 0.97                                                                                                                                                   |
| Grey and Kennedy      | 1993 | 40 spinal cord injured patients                                               | Clinicians rating within 6 weeks pre-discharge, and patient self-rating by self report postal questionnaire at 4 weeks post discharge.                                                                  | r = 0.628 (p<0.0001) r = 0.841 (self care) r = 0.710 (sphincter) r = 0.733 (transfers) r = 0.454 (locomotion) r = 0.029 (communication) r = 0.085 (social cognition) |
| Segal et al           | 1993 | 57 spinal cord injured patients                                               | Two institutions (acute and rehabilitation settings) rating FIM scores, within 6 days of one another. Rating of items was discipline specific. Data gathered retrospectively. | r = 0.83 (all diagnosis) r = 0.87 (complete quadriplegia) r = 0.49 (incomplete quadriplegia) r = 0.74 (complete paraplegia) r = 0.85 (incomplete paraplegia) | r = 0.77 ‘above average’ (feeding) r = 0.46 - 0.67 ‘above average’ (self care) r = 0.6 - 0.72 ‘above average’ (transfers) r = 0.62 ‘above average’ (mobility) r = 0.10 - 0.16 ‘below average’ (sphincter) r = 0.08 - 0.02 ‘below average’ (communication) r = 0.25 - 0.39 ‘below average’ (social cognition) |
| Brosseau and Wolfson  | 1994 | 81 MS patients                                                                | 2 trained physiotherapists scoring each patient independently by interview.                                                                                                                                 | ICC = 0.83                                                                                                                                                  |
| Chau                  | 1994 | 198 rehabilitation inpatients aged between 8-24 years                         | 2 groups of clinicians (therapists and educators) rating the patients most familiar with them.                                                                                                                                 | KF = 0.91                                                                                                                                                   |
| Kidd et al            | 1995b| 25 rehabilitation inpatients                                                 | Patients rated independently by both multidisciplinary team assessment and interview from a single person (motor domain).                                                                             | KF > 0.80 (self care) KF > 0.80 (sphincter) KF > 0.80 (transfers) 0.7 < KF < 0.8 (locomotion) 0.7 < KF < 0.8 (communication) 0.7 < KF < 0.8 (social interaction) KF = 0.63 (memory)    |
| Hobart et al          | 1996d| 64 rehabilitation inpatients                                                 | Patients rated through multidisciplinary team consensus                                                                                                                                              | ICC (motor domain) = 0.97 ICC (cognitive domain) = 0.88                                                                                                                                                               | ICC ranging from 0.70 - 0.97 (values for specific subscales not reported) |

ICC = intraclass correlation coefficient; r = Pearson's correlation coefficient; KF = Fleiss weighted kappa coefficient
Reliability between raters in scoring individual items has also been assessed. Using patient interview as the method of scoring, Brosseau and Wolfson (1994) found the interrater reliability, between two trained physiotherapists, to be good for all motor items \((k = 0.50 - 0.70)\) except locomotion \((k = 0.16)\). In contrast, values for the communication and cognition items were poor \((k = 0.14 - 0.32)\). Using discipline specific scoring as the rating method, Segal and colleagues (1993) found similar results, with the majority of motor items having ‘above average’ reliability coefficients \((r = 0.46 - 0.77)\), and communication and social cognition items having ‘below average’ reliability coefficients \((r = 0.02 - 0.39)\).

Comparisons, in terms of reliability, have been made between other disability instruments. One study found interrater reliability of the motor domain of the FIM to be comparable to the Barthel Index (Kidd et al 1995b), while higher levels of interrater reliability were demonstrated compared to the Uniform Continuous Scale (Chau et al 1994).

**Test-retest reliability:**

Hobart et al (1996d) found high test-retest reliability of the FIM for both the domains and the subscales. He reported intraclass correlation coefficient values of 0.95 for the motor domain, 0.84 for the cognitive domain, and 0.79-0.98 for the subscales.

A number of different methods of scoring the FIM are reported in the literature including: telephone interview (Fuhrer et al 1992); interview (Brosseau and

2) Validity

Validity has been assessed in several ways:

**Content validity**

The task force responsible for the development of the FIM was comprised of clinical, research and administrative experts in the field of rehabilitation, who extensively reviewed the literature before agreeing on an instrument (Granger et al 1986). This provides evidence for its content validity. Content validity was also evaluated by means of specific questions to a variety of rehabilitation professionals regarding: understanding items (88% did not have difficulty); unnecessary items (97% felt this was not the case); and items which should be added (83% felt no need for more items) (Hamilton et al 1987). This adds further evidence of its content validity.

**Construct validity**

A variety of approaches to determine construct validity have been used:

*Convergent construct validity*

This determines the extent to which the FIM correlates with other measures of disability. The motor domain of the FIM has been shown to correlate moderately to highly with: the Disability Rating Scale in head injured patients ($r = 0.641$, $p<0.05$, Hall et al 1993); serial assessments of the Barthel Index in spinal cord injured patients ($0.83 < r < 0.89$, Roth et al 1990) and stroke patients ($0.76 < r < 0.91$, Wagner and Zucchigna 1988); the Quadriplegia Index of Function (feeding $r = 0.75$; bathing $r = 0.92$; grooming $r = 0.94$) and Upper
Extremity Motor Score in quadriplegics (feeding \( r = 0.53 \); bathing \( r = 0.75 \); grooming \( r = 0.91 \), Marino et al 1993).

With particular reference to the MS population, Granger and colleagues (1990) reported the FIM to correlate highly with the Barthel Index, the Incapacity Status Scale, and the Environmental Status Scale, although no correlation figures were provided. Moderate to high correlations were found with the FIM total score and EDSS scores \( (r = -0.91) \), and the FIM locomotion item and EDSS score \( (r = -0.58, \) Brosseau and Wolfson 1994).

In comparison, studies providing evidence for the validity of the cognitive domain are few, and their results are wide ranging and far less convincing. For example, no correlation between any neuropsychologic test results was demonstrated in spinal cord injured patients (Davidoff et al 1990). This is not surprising considering the very low severity and prevalence of significant cognitive deficits in the spinal cord injured population. In contrast, good correlations in head injured patients were shown with the Disability Rating Scale \( (0.728, p < 0.05, \) Hall 1993), and the Orientation Group Monitoring System during the post traumatic amnesia phase \( (r = 0.64, p < 0.001, \) Kaplan and Corrigan 1994). The evidence regarding the validity of the cognitive domain is inconclusive and interpretation of study results should be considered with caution.

With regard to the FIM's ability to assess burden of care, it has been demonstrated that FIM motor domain scores correlate highly with: the need for
help by community care services (as measured in minutes of assistance provided per day) in people with MS (\( r = -0.74 \) to -0.84, Granger et al 1990), stroke (\( r = -0.67 \) to -0.87, Granger et al 1993), and a variety of neurological conditions (\( r = -0.76 \), Disler et al 1993). These support the theoretical premise upon which the FIM is based, that disability measurement can be related to burden of care.

**Group differences construct validity**

This measures the extent to which the FIM is able to detect differences in groups known to differ in terms of disability. Amongst rehabilitation inpatients with a wide range of diagnoses, the total FIM score and its six sub-scales have proven able to detect differences on the basis of: age, co-morbidity and discharge status (Dodds et al 1993; Oczkowski and Barrecca 1993); severity of impairment in spinal cord injury and stroke patients (Dodds et al 1993); severity of head injury (DiScala et al 1992; Kaplan and Corrigan 1994); item difficulty between different impairment groups (Heinemann et al 1993); and patients in different categories of residential accommodation (Disler et al 1993).

**Criterion related validity**

**Predictive validity**

Admission scores on the motor domain, and to a lesser degree on the cognitive domain, have been found to be predictive of rehabilitation length of stay, and discharge function for a wide range of impairment groups (Oczkowski and Barecca 1993; Heinemann et al 1994). Oczkowski and Barecca (1993) also
found that the individual bladder and bowel item admission scores were predictive of discharge location.

Prediction of the patients physical care needs in terms of amount of assistance required from another person has been demonstrated in MS (Granger et al 1990), stroke (Granger et al 1993), and a variety of neurological conditions (Disler et al 1993). By use of multiple regression analysis, Granger and colleagues (1990) compared the ability of the FIM to other commonly used disability instruments (Barthel Index, Incapacity Status Scale, Environmental Status Scale) in predicting the amount of care required by MS patients. They concluded that the FIM motor domain score was the strongest predictor accounting for nearly all of the variance ($r^2 = 0.9982$, $p<.00001$) in help per minutes per day.

3) Responsiveness

A number of studies have shown that the FIM is able to demonstrate statistically significant changes in scores between admission and discharge in a wide variety of medical conditions including: MS (Kidd et al 1995a), stroke (Oczowski and Barrecca 1993; Granger et al 1993), head injury (Hall et al 1993; Cook et al 1994), spinal cord injury, low back pain and orthopaedic conditions (Dodds et al 1993). Hobart and colleagues (1996e) investigated responsiveness of the FIM along with five other disability measures in a sample of 64 MS rehabilitation inpatients. They found that the FIM (and the Barthel Index) were the most responsive of the six instruments (effect sizes 0.32 and 0.39 respectively, $p<0.001$).
Statistical significance however does not assure clinical significance (Johnston and Granger 1994). To date, no information is available to indicate what is considered to be a clinically significant change in FIM scores. It has been shown, however, that a change in total FIM score of one point is equivalent to an average of 2.19 minutes of help from another person per day in people with stroke (Granger et al 1993), 3.38 minutes in MS (Granger et al 1990), and 4.5 minutes with a variety of neurological conditions (Disler et al 1993). This information is useful in interpreting what a change in FIM score may mean in terms of burden of care.

Some authors suggest that the FIM is more sensitive to change than other available disability measures. For example, Bowers and Kofroth (1989) reported it to be 'more sensitive' than the Disability Rating Scale in head injured persons, and Kidd et al (1995b) found it detected change in 84% of patients against 67% on the Barthel Index between admission and discharge from a neurorehabilitation unit. However, Hobart and colleagues (1996e) found little difference between the responsiveness of the Barthel Index and the FIM; and Marino and colleagues (1993) showed that select items on the Quadriplegia Index Function Scale were better able to identify change in quadriplegics than related FIM items.

4) Feasibility

Use of the FIM in routine clinical practice has been demonstrated by a number of centres (e.g. Dodds et al 1993; Peter 1994; Freeman et al 1996a), with estimated times of scoring ranging from between 10-45 minutes (Hall et al
Many agree that training of raters is essential to ensure good reliability (Granger et al 1986; Chau et al 1994; Kidd et al 1995b). Telephone interviews using a standardised format (Fone FIM, 1990) are now being used (Hall 1993; Aisen et al 1996) which, if proven reliable, may make long term follow-up more feasible.

In summary, the FIM motor domain was chosen for use in the studies presented in this thesis, in preference to other instruments. It has proven to be clinically useful, relevant to MS, reliable, valid and responsive in studies evaluating rehabilitation. Evidence in support of the cognitive domain however remains sparse and is conflicting, and its use in assessing neuropsychological needs of specific patient groups has been questioned. The relatively widespread use of the FIM allows comparisons to be made in relation to demographics, severity of disability, and its direction of change with intervention.

2.5 Measuring Handicap

There are many accounts of the profound impact that MS can have on diverse aspects of social activity and family life. As discussed (Section 2.1.2), handicap is the term used to describe these disadvantages. In the context of health experience, it takes into account the physical and psychological effects of disease, as well as relevant non-health factors such as the physical and social environment (Harwood and Ebrahim 1995). Addressing these issues is now considered central to the management of chronic disease (Ebrahim 1990; Segal and Schall 1995; Pearce and Kirchner 1995), and the reduction of
handicap is recognised as a relevant and important goal of rehabilitation (Batavia 1992; BSRM Working Party Report on MS 1993; Harwood et al 1994a).

From the patients point of view handicap issues are the most important consequences of ill-health (Harwood et al 1994a). The wide range of the disadvantages experienced in MS are illustrated by the results of a UK population survey (McLellan et al 1989). Handicaps reported include: premature retirement in 80% of people; difficulty or inability getting out of the house as often as wished by 85% of those severely disabled\(^1\); 97% of severely disabled people were unable to leave the house on a daily basis; 44% had unsuitable transport facilities; 28% had no-one to go out with; 26% were unable to obtain wheelchair access to the places they wanted to go; 25% required personal assistance every day to continue to live at home; and 90% of people regarded themselves as physically handicapped.

While many causes of handicap are beyond the control of health care services (e.g. societal attitudes, economic resources, public transport facilities, and architectural barriers), a proportion of problems can be influenced by rehabilitation intervention. Often these changes occur indirectly and are dependent upon local social services undertaking recommendations made by the rehabilitation team (e.g. aids and adaptations to improve access {ramps, rails, stairlift}, and increase autonomy {wheelchair, environmental controls}). In other instances the changes are directly attributable to rehabilitation intervention (e.g. improved walking may enable access to a wider range of

\(^1\) In McLellan's survey, "severe" disability was defined as an EDSS score of > 7.0.
environments; greater bladder control may increase self esteem and social interaction; the provision of environmental control systems may increase autonomy). Handicap is therefore an important outcome for measuring the effectiveness of intervention.

Other reasons for measuring handicap include:

- its importance and relevance to the patient
- its assessment of the outcome of a variety of interventions by a range of disciplines (e.g. reducing disability, improving motivation, appropriate provision of aids and equipment, rehousing, provision of financial benefits).
  
  This makes it particularly appropriate for evaluating rehabilitation, where the very essence of intervention is an integrated multidisciplinary approach

- in many rehabilitation interventions the impact of the disease is reduced without affecting disability.

2.5.1 Studies in Multiple Sclerosis

To date there has been relatively little research on the application of the measurement of handicap to the area of rehabilitation. It is said to be one of the most poorly measured of all clinical outcomes (Whiteneck et al 1992) with only three rehabilitation studies including it as a variable in their evaluation (Feigenson et al 1981; Kidd et al 1995a; Kidd and Thompson 1997, in press).

These are described in detail in Chapter Three.

Furthermore, while the impact of MS on specific dimensions of handicap (e.g. employment, Rao et al 1991; psychosocial aspects, Perry 1994) has been
reasonably widely reported, very few studies have utilised standardised measurement instruments to describe handicap in more global terms. In 1984, The International Federation of Multiple Sclerosis Societies held a symposium devoted to the use of the Minimal Record of Disability. Three studies reported on handicap issues.

Kuroiwa et al (1984) determined a profile of severity of 75 Japanese MS patients, randomly selected from 24 institutions throughout Japan. He found that most problems occurred in work status and transportation. He used the Environmental Status Scale (ESS) to determine handicap levels. Forty five percent of people experienced problems with work status, 21% required assistance with personal assistance and 31% needed community social services. Correlating ESS scores with FS scores, strong associations were demonstrated with all ESS domains and the pyramidal system (r = 0.73 - 0.75). Personal assistance and transportation domains were strongly associated with sphincter dysfunction (r = 0.73 - 0.76).

De Rham (1984) gathered similar information on a group of 100 patients in Switzerland. The information reported was limited: 71% of people were able to maintain their usual financial standard, although 43% required financial assistance to achieve this; 66% required community health services; and, 16% had to move from home to live in either institutions or hospitals due to difficulties with their care.
A further study by Bauer and colleagues (1984) involved 120 patients attending a long-term continuous care programme. They found that 33% of patients had an unsatisfactory financial situation, 33% were in need of better housing, 66% required help with personal assistance, 33% had an endangered family situation, 66% were unemployed or pensioned, 66% were unable to do housework, and in 33% transportation was possible only in special vehicles.

A decade later, Cervera-Deval and colleagues (1994) reported the level of handicap of 72 MS patients admitted to a Spanish MS Unit. Using the ESS to measure handicap and Kurtzke’s FS to assess impairments, they investigated levels of handicap in relation to neurological impairments. Pyramidal dysfunction was the most frequent neurological finding (86% of patients) and the need for personal care was the most common social handicap (54.7% of patients). Significant relationships were demonstrated between five of the seven ESS items and the pyramidal system (p < 0.05). Cerebellar dysfunction and visual disturbance were significantly associated (p < 0.05) with transportation problems, the need for personal care, and social activity. They concluded that pyramidal dysfunction may be the most important factor to be considered for optimal social integration of these patients.

These studies are valuable in describing the profound impact that MS has on an individuals daily life. They highlight the necessity for measuring handicap in clinical trials.
2.5.2 Instruments for measuring handicap

Handicap is extremely complex and difficult to quantify (Ebrahim 1990). It is a subjective concept, concerned with the value attached to an individual’s situation or experience when it departs from the norm (Wilkin et al 1992). Two people suffering from the same impairment or disability may therefore differ in the type and severity of disadvantage because of differences in expectations and ambitions, and in the extent to which the disability interferes with their normal activity. As a consequence few instruments are currently available to measure this elusive construct (Wade 1992). Fewer still have been comprehensively assessed in terms of their psychometric properties or applied to the rehabilitation setting (Whiteneck et al 1992).

Little agreement exists as how to best measure handicap. Methods include: self-report questionnaires (e.g. the London Handicap Scale, Harwood et al 1994b); visual analogue scales (Wade 1992); semi-structured interviews (e.g. the Environmental Status Scale, Stewart et al 1995); structured interviews (e.g. CHART, Whiteneck et al 1992) or a combination of structured interview, observation and examination (e.g. WHO Handicap Scale, Grimby et al 1993).

The majority of instruments have been developed in reference to the WHO’s system of description (Mellerup 1981; Affleck and Hunter 1988; Whiteneck et al 1992). This system comprises six dimensions: mobility, orientation, occupation, social integration, physical independence, and economic self-sufficiency. While these dimensions are not the only ones possible, it is generally considered that they adequately represent this construct (Ebrahim 1990).
The Environmental Status Scale (ESS) was developed to assess handicap as part of the Minimal Record of Disability for MS (Mellerup et al 1981). It did not however meet the criteria for use as an outcome measure in the studies undertaken in this thesis (refer to Section 2.2). It has been rarely used in studies evaluating clinical outcome (Sharrack and Hughes 1996), and evidence of its psychometric properties is sparse. It has been criticised for a number of issues including: mixing disability and handicap, having limited validity and a misleading scoring system (Stewart et al 1995); being difficult to interpret (de Rham 1984) and; being limited in its scope of enquiry (Battaglia et al 1984). In addition it is an ordinal scale and hence quantitative comparisons are limited. In preference, the London Handicap scale was chosen as the measurement instrument in these studies.

2.5.3 The London Handicap Scale

The London Handicap Scale (LHS) is a generic health status measurement instrument of handicap developed by Harwood and colleagues (1994b). It was specifically designed for use in patients with chronic, multiple or progressive diseases, and in the evaluation of interventions used in their treatment, including rehabilitation.

The LHS is a six item self report questionnaire, based on the dimensions of the WHO's description of handicap (Table 2.8). It emphasises what someone is able to achieve in everyday life, in their normal physical environment, regardless of the help that might be required in the form of human assistance, aids or adaptations. Responses to each question are coded one to six, with one
being the least disadvantaged and six the most disadvantaged, thus forming a
descriptive profile of handicaps for an individual (Appendix 4).

Table 2.8 Dimensions on the London Handicap Scale

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Topics covered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mobility</td>
<td>Ability to get from one place to another, using any help, aids or means of transport normally available</td>
</tr>
<tr>
<td>Physical Independence</td>
<td>Ability to look after self, for example: housework, shopping, looking after money, washing, dressing, toileting.</td>
</tr>
<tr>
<td>Social integration</td>
<td>Describes social contact with other people</td>
</tr>
<tr>
<td>Occupation</td>
<td>What one does with one's time including work, household, and leisure activities</td>
</tr>
<tr>
<td>Orientation</td>
<td>The ability to understand and navigate one's immediate environment, including memory, perception and communication</td>
</tr>
<tr>
<td>Economic self-sufficiency</td>
<td>The effects of ill-health on the ability to earn a living and the ability to use wealth to overcome disadvantages associated with ill-health</td>
</tr>
</tbody>
</table>

Harwood and colleagues (1994a) recognised that quantifying the effectiveness of interventions usually involves comparison of data between different individuals or groups, ideally by the use of scales with interval properties. By using a market research technique, they assigned scale weights to scores on each dimension, calculated from the opinions of 224 subjects from the general population. This enabled scores for each dimension to be assigned a scale weight which can then be combined into an overall handicap score according to a given formula. This total score, with a possible range of zero (maximum possible handicap) to 100 (no handicap), is an estimate of the relative desirability or ‘utility’ of the particular combination of disadvantages (Harwood and Ebrahim 1995). While this scaling technique enables the LHS to be used to make comparisons between groups, it is not valid in comparing individual patients. Comparisons on an individual level can be made by using the
descriptive profile to indicate where particular problems are experienced and where changes have occurred.

2.5.4 Psychometric properties of the LHS

Like virtually all available handicap scales the use of the LHS is in its infancy, and consequently evidence regarding its psychometric properties and application to differing study populations is limited. Studies however are ongoing. It is noteworthy that it has recently been recommended as an approved instrument by the Medical Outcomes Trust, an American organisation specifically established to distribute high-quality standardised health care measurement instruments (Medical Outcomes Trust Bulletin 1996).

1) Reliability

Internal consistency

Moderate to high overall internal consistency has been demonstrated, with Cronbach’s alpha (\(\alpha\)) ranging from 0.67 to 0.88 (Harwood and Ebrahim 1995, Harwood et al 1996). In addition factor analysis has revealed a dominant factor accounting for 55-60% of the variance. This evidence suggests the LHS measures a single unified construct.

Interater reliability

Almost perfect agreement has been demonstrated for the interview technique when gathering LHS data in the general population (n = 644; mean difference in handicap score 0.4/100; limits of agreement 3.5 - -2.7) (Harwood and Ebrahim 1995).
Test-retest reliability

Satisfactory levels of test-retest reliability have been demonstrated in four studies undertaken by Harwood and Ebrahim (1995) (Table 2.9)

Table 2.9 Test retest reliability studies on the LHS

<table>
<thead>
<tr>
<th>Population</th>
<th>Test-retest interim</th>
<th>ICC</th>
<th>Bias</th>
<th>LOA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stroke (n=37)</td>
<td>2 weeks</td>
<td>0.91</td>
<td>+0.5</td>
<td>+/- 17</td>
</tr>
<tr>
<td>Rheumatology (n=79)</td>
<td>3 months</td>
<td>0.93</td>
<td>+0.9</td>
<td>+/- 20</td>
</tr>
<tr>
<td>Intervention studies (n=16)</td>
<td>?</td>
<td>0.77</td>
<td>+0.1</td>
<td>+/- 17</td>
</tr>
<tr>
<td>General population (n=50)</td>
<td>17 weeks (mean)</td>
<td>0.84</td>
<td>-6.4</td>
<td>+/- 21</td>
</tr>
</tbody>
</table>

ICC = Intra class correlation coefficient; Bias = mean difference between test and retest scores out of 100; LOA = Limits of agreement at 95% confidence interval; - = more handicapped, + = less handicapped.

2) Validity

Content validity

The six dimensions of the LHS conform to the WHO classification system, which is generally agreed to provide a comprehensive profile of handicap. Both lay subjects and health care professionals were involved in reviewing and piloting the LHS questionnaire to ensure it remained simple and unambiguous, whilst remaining faithful to the WHO classification categories (Harwood and Ebrahim 1995). These factors provide evidence for the content validity of this scale. With regard to the validity of the scale weighting, it was demonstrated that the ordering of scores within each dimension, and the relative contribution of each level to overall desirability, followed the expected pattern (Harwood and Ebrahim 1995).
**Criterion related validity**

**Predictive validity**

Excellent predictive validity, when compared with direct estimates of severity made by subjects, has been demonstrated ($r = 0.98$, $p<0.0001$, Harwood et al 1994b,c).

**Construct validity**

**Convergent construct validity**

The LHS has been shown to correlate moderately well, or higher, with other scales measuring handicap or closely related concepts. Examples include correlation with: the modified Rankin Scale in patients six months post-stroke ($r=0.78$); all dimensions of the Disease Repurcussion Profile in rheumatoid arthritis ($r^2 = 0.13 - 0.46$, $p<0.0005$); and the CARE Activity Limitation Scale ($r = 0.85$), and interference with life question ($r = 0.67$) in community patients (Harwood and Ebrahim 1995).

**Discriminant construct validity**

The LHS has shown to correlate less well with health status scales not measuring handicap (Harwood and Ebrahim 1995, Harwood et al 1996) (Table 2.10). A range in the strength of correlations between the LHS and other scales is observed. A weak association is seen with diagnostic measures, moderate association with measures of impairment, and a stronger degree of association with measures of disability. Harwood and Ebrahim (1995) state that these relationships are in the directions expected, and the grading of associations is plausible.
Table 2.10 Associations between LHS and other health status measurements

<table>
<thead>
<tr>
<th>Population</th>
<th>Scales and strength of association (r²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post stroke</td>
<td>Barthel Index (0.31), Nottingham Extended ADL (0.48), Nottingham Health Profile (0.27), Geriatric Depression Score (0.18), Guy's Stroke Prognostic score (0.15)</td>
</tr>
<tr>
<td>Arthritis clinic</td>
<td>Stanford Health Assessment Questionnaire (0.58)</td>
</tr>
<tr>
<td>Community survey</td>
<td>Activities of Daily Living Scale (0.53); OPCS disability categories (0.72); Depression scale (quintiles) (0.18); Dementia scale (quintiles) (0.43); Grip strength quartiles (0.2); Visual impairment (0.18); Total number of impairments (0.37); Total number of diagnoses (0.17)</td>
</tr>
<tr>
<td>Arthroplasty candidates</td>
<td>Extended ADL scale (0.36); SF-36 energy/vitality (0.16); SF-36 health perceptions (0.2); SF-36 mental health (0.11); SF-36 pain (0.08); SF-36 physical functioning (0.18); SF-36 social functioning (0.34)</td>
</tr>
<tr>
<td>Geriatric day patients</td>
<td>Barthel index (0.26)</td>
</tr>
</tbody>
</table>

r² = the proportion of handicap variance explained by variability in the comparator scale

Comparison of different populations demonstrate that grading of severity, determined by the overall score on the LHS, is in line with clinical experience with community residents having the highest score (i.e. at least disadvantage) and stroke patients the lowest scores (Harwood and Ebrahim 1995).

**Hypothesis testing**

Harwood et al (1994b) demonstrated that, as hypothesised, the level of handicap assessed by the LHS, was greater in those people with higher levels of disability, worse perceived health status, greater depression, and lower satisfaction with life.
3) Responsiveness

The LHS is capable of demonstrating changes in a variety of interventions including total hip and knee replacements, coronary angioplasty, attendance at a geriatric day hospital, and late recovery after stroke (range of effect sizes 0.07 - 0.85, Harwood and Ebrahim 1995). This demonstrates the ability to detect changes ranging from small to large. These changes are in line with clinical experience, with the largest improvement seen in hip replacement patients, a moderate effect size for angioplasty, and a small effect size for geriatric day patients. The developers of the scale, however, caution in the interpretation of this data since the estimates are imprecise (as evidenced by wide 95% confidence intervals) and data is uncontrolled. Despite these limitations, they claim the scale is ‘acceptably responsive for a generic scale’.

4) Feasibility

The LHS questionnaire is short, taking a maximum of 10 minutes to complete. It has shown to be easily understood and acceptable to patients (Harwood et al 1994b). A comprehensive operational and scoring manual, with detailed interpretation guidelines is available (Harwood and Ebrahim 1995).

In summary, choosing an instrument to measure handicap was difficult. Few instruments have been developed and even fewer have been well evaluated. The LHS was chosen for use in the studies presented in this thesis in preference to other instruments. It is relevant to the MS inpatient rehabilitation population, clinically useful, and demonstrates scientifically sound evidence of
its psychometric properties in neurological patients and those with chronic
disease.

2.6 Measuring Health-related Quality of Life

Extending beyond the disease-handicap continuum of the ICIDH is the concept
of quality of life (Figure 2.1). This concept broadens the framework of the
disablement model to view health in more positive terms as "a state of complete
physical, mental and social well-being and not merely the absence of disease
or infirmity" (WHO 1948 cited by Führer 1994).

Figure 2.1 A conceptual and operational model for measuring outcome

<table>
<thead>
<tr>
<th>OUTCOME</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONCEPTUAL CONTINUUM</td>
</tr>
<tr>
<td>Disease</td>
</tr>
<tr>
<td>OPERATIONAL CONTINUUM</td>
</tr>
<tr>
<td>MS</td>
</tr>
<tr>
<td>Ataxia</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

(reproduced with permission from the author, Tennant 1996).

An estimated 40% of the total health service budget is expended on treatments,
such as rehabilitation, in which the primary aim is to impact on the quality
rather than the quantity of a person's life (Hopkins 1992). As a consequence
the measurement of quality of life (QOL) is increasingly recognised as relevant
in assessing the outcomes of medical care (Rodriguez et al 1994; Devinsky
1995), and is considered by some to be central to the evaluation process (Ware 1993a; Johnson 1996).

2.6.1 Defining “Health related quality of life”

No single universally accepted definition of QOL exists (Stensman 1985; Spilker 1990; Pfenning et al 1995). This has led to confusion about the meaning and use of terms such as 'general well-being', 'subjective health status', 'functional status', 'quality of life' and 'health-related quality of life' which are frequently used in the literature in an apparently equivalent manner (Bowling 1991; Fitzpatrick et al 1992b). It has therefore been suggested that researchers define the conceptual issues being investigated in order to clarify what is being measured in QOL studies (Selby 1992). The literature does not reflect this. In a critical appraisal of QOL measures, Gill and Feinstein (1994) noted that of the 75 articles reviewed only 15% of the investigators defined QOL, and only 35% gave reasons for selecting the chosen measures.

Health care researchers are most interested in aspects of life closely related to health status. The dimensions of QOL, which are most directly affected by disease and health care intervention, are termed ‘health related quality of life (HRQOL)’. This has been defined by Nelson (1990, cited in Hopkins 1992) as “the subjective satisfaction with health-related dimensions of life as judged by the patient himself”, and is the construct measured in this thesis. This enables an evaluation of the patients point of view of how they perceive and react to their disease (Hobart et al 1996a).
2.6.2 Conceptual issues

Although difficulties continue to exist with conceptual, philosophical and ethical issues concerning HRQOL, there now appears to be widespread agreement on some fundamental aspects. These include:

i) QOL as an all-embracing concept

QOL embraces and is influenced by a myriad of closely inter-related factors including: basic standards of living and security, psychological and physical factors, and relations to other people and society (Athlsio et al 1984; Stensman 1985). Some suggest that good health is one of the primary requisites to the enjoyment of a high QOL and hence should be added to this baseline of needs (Fallowfield 1990). Indeed in a study examining the most valued end states of existence, Rokeach (1973), found that every subject valued good health higher than any other option (cited in Fallowfield 1990).

ii) The multi-dimensional nature of health related QOL.

HRQOL is a multi-faceted phenomenon (Bowling 1991). General agreement exists as to the core set of health dimensions contributing to this global concept (e.g. Fallowfield 1990; Spilker 1990; Walker and Rosser 1992). These include:

- physical health and functional abilities (e.g. pain, mobility, sleep)
- psychological status and well-being (e.g. depression, adjustment to illness)
- social functioning (e.g. personal relationships, leisure activities)
- role functioning (e.g. paid or unpaid employment, household tasks)
- economic status (e.g. financial status)
The measurement of QOL must address both the variety and the inter-relationships between these core dimensions (Patrick and Bergner 1988; Bowling 1991).

iii) The variation in individuals' perception of QOL

Just as handicap is inherently subjective, so too quality of life means different things to different people, reflecting a wide array of knowledge, experiences, and values (Spilker 1990). It is highly dependent upon the individual's perception of their own needs and expectations and is rated depending upon what is seen as the frame of reasonable or attainable alternatives (Minair 1991). This reference range differs between different group's, for example the young and the old, the physically and non-physically disabled (Evans 1992).

The way in which a disease impacts on a person's life is not only unique to the individual but varies over the course of their lifetime. QOL is therefore a relative concept that may change in relation to the individual's needs and expectations at a specific point in time. This is particularly important to consider in MS, where the disease course spans many decades, the pattern of disease is variable and unpredictable, and where symptoms may fluctuate on an almost daily basis.

2.6.3 Measurement issues

The pre-requisite for measuring a concept is a clear definition. As discussed, this is very difficult with a concept as multi-faceted, subjective and individual as HRQOL. Consequently both its measurement and interpretation of results has
proven extremely complex (Lankhorst 1989), with some questioning whether it is even open to quantitative assessment (e.g. Bergsma and Engel 1988; Wade 1992). It is recognised however that the analysis (definition, measurement and evaluation) of HRQOL is in its early stages of development (Dossa 1989), with continued advances being made (Pfennings et al 1997, in press).

Models of measurement

Two distinct models of measurement are available to measure HRQOL:

• the generic model which is designed to be broadly applicable across different types and severity's of disease, medical interventions and demographic and cultural groups, and thereby allows comparison across different population's

• the condition-specific model which is designed to reflect clinically relevant symptoms for a specific disease, and is not concerned with establishing global standards (Bowling 1995).

It is recommended that studies include both a condition-specific and a generic instrument (Patrick and Deyo 1989; McHorney et al 1992; Garratt et al 1993).

2.6.4 Studies in Multiple Sclerosis

The concept of measuring quality of life has been slow to transfer to the areas of neurology and rehabilitation (Gill and Feinstein 1994; Devinsky 1995; Bowling 1995). Only one study has evaluated it as an outcome of MS inpatient rehabilitation (Jonsson et al 1996). Few have systematically examined it in the MS population (Rudick et al 1992; Borgel et al 1992; Brunet et al 1996; Hermann et al 1996; Freeman et 1996b; Aaronson 1997). The majority of these
studies have used a different instrument, or battery of instruments, to investigate HRQOL, making comparison of results difficult. Furthermore, in a number of the studies the generalisability of findings is limited by the relatively small and select samples. For example, the samples were comprised of rehabilitation inpatients (Borgel et al 1992; Freeman et al 1996b) and included only those with a disease duration of longer than 10 years (Rudick et al 1992; Borgel et al 1992).

Using a standardised interview (The Farmers QOL Index), Rudick and colleagues (1992) investigated the HRQOL of patients with three chronic diseases - Inflammatory Bowel Disease (IBD), Rheumatoid Arthritis (RA), and MS. The results suggested that HRQOL was best in the IBD group and worst in the MS group. Significant differences were demonstrated between each disease (p<0.017), suggesting that unique clinical profiles differentially characterised each of the diseases. All MS patients were also assessed with Kurtzke's EDSS and FS (EDSS: ≤ 5 in 50%; 6.0 - 6.5 in 15%; ≥ 7.0 in 35% of patients). No correlation was found between the total QOL score and EDSS score (r = -0.28, p<0.02). Of the FS scales, only the Visual Scale score correlated significantly with the QOL total (r = -0.5, p<0.0002) and subscale scores (r = -0.39 - -0.43, p<0.005) suggesting that vision is related to QOL in MS. The duration of MS appeared to be unrelated to QOL scores (p < 0.0329).

Borgel and colleagues (1992) reported the preliminary findings of a study of 35 MS rehabilitation inpatients. Using assessments of physical disability (Barthel Index and EDSS), depression (Montgomery and Asberg's Depression Rating
Scale) and self-perceived health status (The Nottingham Health Profile), he compared them with 47 stroke patients. Results demonstrated that: physical disability was greater in the MS population (mean EDDS 6.0, p<0.0003); mild to moderate depression was equivalent in both groups (53% of patients); and perceived morbidity in terms of HRQOL was higher in the MS group particularly in the areas of energy (p = 0.02), physical mobility (p = 0.02), and pain (p = 0.02). An increase in physical disability (EDSS>6.0) was associated with an increase in depression (r² = 0.455, p=0.0001) and morbidity in perceived health status (r² = 0.209, p = 0.003). This contrasted with the results by Rudick and colleagues (1992).

In the past year, three studies have reported the use of the generic quality of life measurement instrument, the 36-item Short Form Health Survey Questionnaire (SF-36) in people with MS (Hermann et al 1996; Brunet et al 1996; Freeman et al 1996b). The former compared the HRQOL of three chronic disease groups: epilepsy (n=271), MS (n=85), and diabetes (n=555). The information was collected by mail survey. Patients in all groups represented a spectrum of severity from the newly diagnosed to those with chronic and severe problems. The results demonstrated that patients with MS reported significantly worse HRQOL than both the epilepsy and diabetes groups (who did not differ from one another) in the dimensions of physical functioning, physical role limitations, energy/vitality and social function. With respect to emotional dimensions (emotional well-being and role limitations), patients with MS and epilepsy did not differ from one another, but both reported significantly worse HRQOL than the diabetes group.
The latter two studies each used a prospective study design to investigate HRQOL using the SF-36. Brunet and colleagues (1996) gathered information from 97 patients attending the MS clinic of a District General Hospital. Like Hermann (1996), they found that the patients reported particularly poor HRQOL in three dimensions: physical functioning, physical role limitations and energy/vitality, scoring well below those of a representative sample of the general US population. Freeman and colleagues (1996b) investigated the HRQOL of 50 MS rehabilitation inpatients, all of whom were moderately to severely disabled (mean EDSS 6.5, range 6.0 - 9.0). Similar to Brunet and colleagues (1996), the SF-36 scores in their population were compared to normative values for the general UK population. The results were in agreement with those of the previous studies - poorer levels of HRQOL were reported by the MS patients in all SF-36 dimensions compared to the general healthy population, particularly in the dimensions of physical functioning and role functioning.

Recently, Aaronson (1997) conducted a large, population-based, postal survey of persons with MS and their caregivers (n = 697 and 345 respectively). By using the same QOL questions as those in the General Social Survey (a Canadian Social Survey Questionnaire), he was able to compare his sample with the Canadian general population (1692 disabled and able-bodied persons of the same sex, age and education level). Results demonstrated that MS persons reported less satisfaction with several QOL components compared to the disabled general population (general health, occupation, life as a whole).
Carers also reported less satisfaction compared to the general able-bodied population (finances, life as a whole). Poorer QOL for MS persons was associated with unemployment, increased severity of symptoms, mobility limitations, an unstable disease course and interference with social activities. In the carers, a poorer QOL was associated with longer duration of caregiving and an increased severity of the symptoms or an unstable disease course of the care recipient.

The only study to have measured HRQOL as an outcome of MS inpatient rehabilitation was recently undertaken by Jonsson and colleagues (1996). Twenty one patients were assessed with the Disability and Impact Profile and the Beck Depression Inventory on admission and discharge. All patients were also rated on the EDSS (mean 6.6, range 3.5-8.0). Following rehabilitation (length of stay ranging from five to eight weeks) the most significant improvements were measured on the behavioural and social aspects of HRQOL (e.g. fatigue \(p<0.01\), mood \(p<0.03\), pessimism \(p<0.04\), and irritability \(p<0.04\)). Improvements in physical aspects were noted in the ability to climb stairs \(p<0.04\) and to ‘do work’ \(p<0.01\).

These studies consistently demonstrate the profound impact that MS has on an individual's daily life. They highlight the need for further research in this area, and the necessity for measuring HRQOL in clinical trials.
2.6.5 Instruments for measuring HRQOL

In the past decade a number of generic instruments have become available for measuring HRQOL (e.g. the Nottingham Health Profile, Hunt et al 1985; Short-Form 36 Health Survey Questionnaire, Ware et al 1993b; the Q-TWiST, Schwartz et al 1995; the Disability and Impact Profile, Lankhorst et al 1996). Many however, lack a clear theoretical basis (Gill 1995). Very recently, three condition-specific instruments have been developed for MS; the Multiple Sclerosis Quality of Life-54 (Vickrey et al 1995), the Functional Assessment of Multiple Sclerosis (Cella et al 1996), and the Leeds MSQoL Scale (Ford et al 1997). Preliminary data for each of these instruments is available, however they are yet to be evaluated in accordance with recommended guidelines (Medical Outcomes Trust 1995). All look promising as useful tools for future outcome studies in MS. At the commencement of this thesis no condition-specific instruments were available. The choice of instrument was therefore limited to a generic measure.

2.6.6 The 36 item Short-Form Health Survey Questionnaire

The SF-36 is currently accepted as the gold standard generic measure of HRQOL (Garratt et al 1993; Pfennings et al 1997, in press). It is the most recent in a series of measurement instruments developed in the United States by the Medical Outcomes Study as a part of an extensive programme of evaluation into HRQOL measures. Developed in 1988-9, it is 36 item questionnaire which has been adapted from longer pre-existing instruments (Appendix 5).
The SF-36 is a self-report measure of health perception. It was designed for use either as a questionnaire or for administration by a trained interviewer in person or by telephone. It is available in several languages and has been used world-wide in clinical practice and research, health policy evaluations, and general population surveys (Ware 1993b). A standard UK version has been developed (Brazier et al 1992). Average scale values, including means and standard deviations, for the UK population (Jenkinson et al 1993) and USA population (Ware et al 1993b, 1994a) are available for purposes of comparison.

The conceptual framework of the SF-36 is based on the multi-dimensional model of health and the belief that the patient’s point of view is central in monitoring medical care outcomes (Ware 1993a ; Brazier et al 1992). It measures eight health concepts thought to be representative of and relevant to everyone’s functional status and well-being. The health concepts chosen represent those most frequently included in health surveys, along with two additional concepts (bodily pain and vitality) which were strongly supported by empirical study (Ware and Sherbourne 1992). Table 2.11 details these conceptual domains and the items measured in them.

The eight separate scales of the SF-36 allow a profile of health status scores to be determined. These scales are scored using Likert’s method of summated ratings (Likert 1932 cited by Ware et al 1993b). For each dimension, item scores are coded, summed, and transformed on a scale from zero (worst possible health state) to 100 (best possible health state). Five of the scales (PF, RP, BP, SF, and RE) define health status as the absence of limitation or
disability. For these scales, the highest possible score of 100 is achieved when no limitations are observed. Three of the scales (GH, VT, and MH) measure a wider range of negative and positive health states. Here a mid-range score is achieved when respondents report no limitations, and a maximum score of 100 is only achieved when health is evaluated positively.

Table 2.11 Conceptual domains of the SF-36

<table>
<thead>
<tr>
<th>Domain</th>
<th>No. of items</th>
<th>Topics covered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-reported general health (GH)</td>
<td>5</td>
<td>Rating of own health, Comparison with other peoples health, Comparison with other's proneness to illness.</td>
</tr>
<tr>
<td>Physical functioning (PF)</td>
<td>10</td>
<td>Extent to which health limits 10 levels/types of physical activity; 3 items refer to walking, 2 ..to climbing stairs, 2 ..to bending and stretching, 1 ..to combined movements.</td>
</tr>
<tr>
<td>Mental Health (MH)</td>
<td>5</td>
<td>Degree of nervousness..calmness, Degree of happiness...sadness.</td>
</tr>
<tr>
<td>Role limitations - physical (RP)</td>
<td>4</td>
<td>Limits that physical health puts on range and extent of all types of work.</td>
</tr>
<tr>
<td>Role limitations- emotional (RE)</td>
<td>3</td>
<td>Limits that &quot;emotional problems&quot; put on range and extent of all types of work.</td>
</tr>
<tr>
<td>Bodily pain (BP)</td>
<td>2</td>
<td>Severity of pain, Impact on activities.</td>
</tr>
<tr>
<td>Energy/ tiredness (VT)</td>
<td>4</td>
<td>How energetic, How tired.</td>
</tr>
<tr>
<td>Social functioning (SF)</td>
<td>2</td>
<td>Impact of physical health or emotional problems on &quot;normal&quot; social activities.</td>
</tr>
<tr>
<td>Change in health in past year (regarded as a supplementary question, not a core domain)</td>
<td>1</td>
<td>Single unscored item comparing health now with 12 months ago.</td>
</tr>
</tbody>
</table>

(Table reproduced from Outcomes Briefing Issue 4, 1994)
The scores on these eight separate scales can be reduced to two summary scores, a physical component and a mental component, by means of principal components analysis (Ware 1994a). The physical component summary score is an index of the physical functioning, physical role, pain and general health scales. The mental component summary scale is an index of the vitality, mental health, role emotional and social functioning scales. The advantage of these component scales is that fewer statistical comparisons need to be undertaken when analysing results, thereby reducing the role of chance in testing experimental hypotheses (Ware et al 1995).

2.6.7 Psychometric properties of the SF-36

The psychometric properties of the SF-36 have been extensively evaluated.

1) Reliability

Internal consistency

High overall internal consistency has been demonstrated by all of the eight scales (α values ranging from 0.78 - 0.93) (McHorney et al 1992; Ware 1993b), and both the physical (α 0.92) and mental (α0.91) component summary scales (Ware et al 1995).

Test-retest reliability

Excellent results at two weeks are reported for the use of the SF-36 in primary care settings in the UK (test-retest reliability coefficient = 0.6-0.81, Brazier et al 1992) and in the USA (McHorney et al 1992). The statistical interpretation of
these results however have been disputed (Hunt and McKenna 1992; Gompertz et al 1992).

2) Validity

Content validity

The SF-36 has content validity. It is based on the health concepts most frequently measured in other widely used survey forms (Ware et al 1993b), and adheres to the standards published by Ware (1987) for the evaluation of content validity of general health measures.

Construct Validity

Convergent construct validity

A number of studies have shown that the SF-36 correlates highly with other available measures of the same criterion (Table 2.12).

Table 2.12 Associations between SF-36 and other health status instruments

<table>
<thead>
<tr>
<th>SF-36 dimension</th>
<th>Scales and strength of association (r)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mental health</td>
<td>Nottingham Health Profile (0.67), Brazier et al 1992</td>
</tr>
<tr>
<td></td>
<td>Sickness Impact Profile-psychological dimension (-0.70), Katz et al 1992</td>
</tr>
<tr>
<td>Physical functioning</td>
<td>Functional Status Questionnaire (0.73), Katz et al 1992</td>
</tr>
<tr>
<td></td>
<td>Physical mobility on the Nottingham Health Profile (-0.52), Brazier et al 1992</td>
</tr>
<tr>
<td></td>
<td>Physical mobility of the Sickness Impact Profile (-0.78), Brazier et al 1992</td>
</tr>
<tr>
<td>Energy/Vitality</td>
<td>Energy scale of the Nottingham Health Profile (0.68), Brazier et al 1992</td>
</tr>
<tr>
<td>Pain</td>
<td>Condition specific Knee Pain Scale (0.41), Katz et al 1992</td>
</tr>
<tr>
<td></td>
<td>Pain on the Nottingham Health Profile (-0.55), Brazier et al 1992</td>
</tr>
<tr>
<td>Social function</td>
<td>Social isolation on the Nottingham Health Profile (-0.41), Brazier et al 1992</td>
</tr>
<tr>
<td>General health</td>
<td>Quality of Well-being Scale (0.52), Fryback et al 1993 (as cited in Ware et al 1993b)</td>
</tr>
</tbody>
</table>
**Discriminant construct validity**

Dimensions considered to be unrelated have been shown to produce differing results. Examples include poor correlations with: the SF-36 physical function dimension and the social isolation on the NHP \( r = -0.2 \), Brazier et al 1992; Katz Activities of Daily Living Scale and the SF-36 mental health domain \( r = 0.10 \), Reuben and Siu 1990 cited by Ware et al 1993b).

**Group differences construct validity**

The SF-36 has the ability to distinguish between groups with expected health differences including: the elderly, people from different socio-economic classes, people with chronic physical problems, and use of hospital and General Practitioner services (Brazier et al 1992; Jenkinson et al 1993; Garratt et al 1993); different medical conditions (Garratt et al 1993); medical and psychiatric conditions (Sherbourne et al 1992 cited by Ware et al 1993b).

These estimates of reliability and validity have been determined from cross sectional studies. It is not yet known whether they will hold true for longitudinal studies. Ware and colleagues (1993b) state however, that according to psychometric theory, measures that discriminate well among different groups of patients at one particular time (as demonstrated by the SF-36) tend to do well in capturing changes over time.

**3) Responsiveness**

Evidence regarding responsiveness of the SF-36 is limited (Garratt et al 1993) and remains an area of contention (Harries and Hill 1994; Mawson 1995).
While different dimensions of the SF-36 are responsive to change in some populations (e.g. the vitality scale in clinical trials of hypertension {Croog et al 1986, cited in Ware et al 1993b} and prostate disease {Fowler et al 1988}), others provide evidence to the contrary (e.g. evaluation of community-based services for older people {Hill and Harries 1994}).

4) Feasibility

The SF-36 has been shown to be easy to use and acceptable to patients in the UK (Brazier et al 1992; Garratt et al 1993) and USA (Ware 1993b), taking approximately 5-10 minutes to complete. High postal response rates have been reported (Brazier et al 1992; Jenkinson et al 1993; Garratt et al 1993). A comprehensive operational and scoring manual, with detailed interpretation guidelines is available (Ware 1993b).

In summary, the SF-36 was chosen for use in the studies presented in this thesis in preference to other measurement instruments. It is the gold-standard generic measure of HRQOL, having been extensively evaluated in terms of both its clinical utility and its psychometric properties in a wide variety of different populations and settings.

2.7 Measuring well-being and mood

This chapter has highlighted the profound economic, social, and interpersonal consequences, and the overall unpredictability and uncertainty that MS presents to both the person and their family. Considering this myriad of
problems, together with the fact that there is no known effective treatment, it is not surprising that depression in MS is common (Larcombe and Wilson 1984; Minden and Schiffer 1991; Sadovnick et al 1996).

The precise prevalence of depression in MS is not known. No large, population-based studies have been undertaken, with estimates determined by small, highly selected samples. The estimates of current prevalence range from 27% (Surridge 1969) to 68% (DePaolo et al 1980). The lifetime prevalence rates are up to three times higher than that of the general population and range from 37-54% (Minden and Schiffer 1991). These variations may be due to differences in sample characteristics, diagnostic techniques, or may reflect the inherent variability of MS.

Typically depression in MS is moderate to severe (Joffe et al 1987; Whitlock and Siskind 1980; De Paulo et al 1980), and has a significant adverse effect on patient's functioning and on their family and social lives (Minden and Schiffer 1990). Mood disturbances such as anger, worry, irritability, discouragement and depressive symptoms are the commonest features evidenced (Whitlock and Siskind 1980; Schiffer and Babigian 1984; Ron and Logsdail 1989). Precise suicide rates are unknown, although it is generally accepted that people with MS are at a heightened risk of suicide (Sadovnick et al 1996).

Clinically there is a general acceptance (as yet unproved) that emotional disturbances such as poor motivation, depression and anxiety may influence the outcome of many health care interventions (Pimm and Tyerman 1991;
This was highlighted by a survey which showed that 82.2\% of MS practitioners believed that depression contributes 'moderately or significantly' to overall disability in MS (Fischer et al 1994). This emphasises the importance of measuring these dimensions, particularly in interventions such as rehabilitation, where a key goal is to improve the patients well-being (Ketelaer and Battaglia 1991) and facilitate their coping processes and adjustment (La Rocca et al 1993).

2.7.1 Studies in multiple sclerosis

Research on the psychological aspects of MS has increased significantly in the past decade. A number of studies have looked at the relationships between emotional distress and disease activity (e.g. Dalos et al 1983; Rabins et al 1986; Mindens and Schiffer 1991). Few, however, have focused on the implications of this research for clinical practice.

Emotional disturbance has been shown to be associated with increasing disease activity (Dalos et al 1983). Dalos and colleagues (1983) were the first to investigate this. In a prospective longitudinal study, patients completed the 28 item General Health Questionnaire on a monthly basis for one year. They found a significant increase in severity as well as prevalence of emotional disturbance during times of progression or exacerbation (increased disease activity) with 39\% of patients in the remission group having abnormal scores compared to 91\% in the exacerbation group (p < 0.001). A further study by McIlvor and colleagues (1984) found that patients, with primarily spinal disease, in the chronic progressive phase were more depressed than those in the
relapsing remitting phase of MS. They did not differentiate however between those in remission and in relapse, making interpretation of their results somewhat difficult.

The relation of patients emotional status to their physical progress is not well understood (Langdon 1996). Some studies provide evidence of a relationship between severity of disability and depression (Whitlock and Siskind 1980; McIlvor et al 1984), while others do not (Dalos et al 1983; Rabins et al 1986; Joffe et al 1987; Schnek et al 1995; Smith and Young 1996). It has been suggested that this disagreement may reflect sample differences such as the percentage of patients in remission or exacerbation, or of those with primarily cerebral versus spinal disease (Minden and Schiffer 1991).

Evidence concerning the relationship between depression and neurological features in MS vary. Some studies suggest that emotional disturbance is related to severity of MRI abnormalities (Horner et al 1987; Reischies et al 1988) while others provide evidence to the contrary (Ron and Logsdail 1989). Features such as types of impairment (e.g. motor, sensory, visual, cognitive) appear unrelated to emotional disturbance (Joffe et al 1987; Devins and Seland 1987; Ron and Logsdail 1989; Mindens and Schiffer 1991; Moller et al 1994). Results of studies on adjustment suggest that a relationship exists between duration of disease and emotional disturbance. Those with a longer duration of disease appear to gradually adapt to their illness over time, displaying fewer coping difficulties (Brooks and Matson 1982; Schnek et al 1995).
On searching the literature, no studies were found which used standardised instruments to investigate the impact that emotional disturbance has on MS rehabilitation outcome. Only one study investigated whether the multidisciplinary inpatient programme was associated with changes in level of emotional distress (Jonnson et al 1996, refer to Section 2.6.4).

### 2.7.2 Instruments for measuring well-being and mood

Most instruments for measuring emotional disturbances have been developed in the field of psychiatry, with very few designed for use with physically ill or disabled people (Wade 1992; Nuyenhuis et al 1995). Methods of assessment include self-report questionnaires (e.g. the General Health Questionnaire, Goldberg 1972; the Hospital Anxiety and Depression Questionnaire, Zigmond and Snaith 1983; the Beck Depression Inventory, Beck et al 1961) and structured interviews. It is generally agreed that structured interviews are preferable to self-report measures (Minden and Schiffer 1990, 1991). They are however more costly, require skilled interviewers, considerable interview training, are time-consuming, and cannot be used to quantify change over time. Of considerable relevance to the studies presented in this thesis, they are of little use in making comparisons between samples. A self-report scale, the General Health Questionnaire (GHQ) was therefore chosen for use in these studies.
2.7.3 The General Health Questionnaire

The GHQ is the most widely applied measure of psychiatric disturbance in the UK (Bowling 1995). It is a self-report instrument designed to detect changes in emotional state (Rabins and Brooks 1981). The GHQ consists of a series of questions about current symptoms, abnormal feelings and thoughts, and aspects of observable behaviour. It is based on the principle that psychological distress depends on a critical number of key symptoms, rather than on one particular symptom (Wade 1992).

The GHQ was designed for use in general population surveys, primary medical care settings and general medical outpatients. It has been used in neurological patients (De Paulo et al 1980; Lykouras et al 1996) and, more specifically, in the MS population (Rabins and Brooks 1981; Dalos et al 1983; Rabins et al 1986; Ron and Logsdail 1989). Within these settings it has been used as a screening instrument, and as both an estimator and indicator of morbidity (Tarnopolsky et al 1979).

Several versions of the GHQ are available. These are known by the number of items in the questionnaire (GHQ-60, GHQ-30, GHQ-28, GHQ-20 GHQ-12). All versions may be used to indicate the presence and the severity of emotional disturbance. Each version covers four identifiable elements of distress: depression, anxiety, social impairment, and hypochondriasis (Table 2.13). Although these separate types of distress are covered, the GHQ is not intended to distinguish among psychiatric disorders or to be used in making diagnoses (McDowell and Newell 1987).
### Table 2.13 Contents of the 28-item GHQ

<table>
<thead>
<tr>
<th>Scale A</th>
<th>Scale B</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Somatic Symptoms</strong></td>
<td><strong>Anxiety and Insomnia</strong></td>
</tr>
<tr>
<td>feeling perfectly well</td>
<td>lost much sleep over worry</td>
</tr>
<tr>
<td>in need of a good tonic</td>
<td>difficulty staying asleep</td>
</tr>
<tr>
<td>run down</td>
<td>constantly under strain</td>
</tr>
<tr>
<td>felt that you are ill</td>
<td>edgy and bad tempered</td>
</tr>
<tr>
<td>pains in your head</td>
<td>scared or panicky</td>
</tr>
<tr>
<td>pressure in your head</td>
<td>everything on top of you</td>
</tr>
<tr>
<td>hot or cold spells</td>
<td>nervous and strung up</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scale C</th>
<th>Scale D</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Social Dysfunction</strong></td>
<td><strong>Severe Depression</strong></td>
</tr>
<tr>
<td>busy and occupied</td>
<td>thinking of yourself as hopeless</td>
</tr>
<tr>
<td>taking longer over things</td>
<td>life entirely hopeless</td>
</tr>
<tr>
<td>doing things well</td>
<td>life not worth living</td>
</tr>
<tr>
<td>satisfied with carrying out task</td>
<td>make away with yourself</td>
</tr>
<tr>
<td>playing a useful part</td>
<td>nerves too bad</td>
</tr>
<tr>
<td>capable of making decisions</td>
<td>dead and away from it all</td>
</tr>
<tr>
<td>enjoy normal activities</td>
<td>idea of taking your life</td>
</tr>
</tbody>
</table>

Items are rated by a two point system, where symptoms are scored as either present (one point) or absent (zero points). The total score and sub-scale scores are added and compared with pre-established cut-off scores. The higher the score, the more distressed the respondent and the higher the probability of the presence of emotional disturbance. These scores can be used to compare two populations either by (1) assessing their proportion of high scores or by (2) comparing their distributions, or mean scores and standard deviations of the scores in the samples (Goldberg 1972; Tarnopolsky et al 1979; Goldberg and Williams 1988).
The wording and scoring of the GHQ questionnaire distinguishes between chronic stable complaints and recent exacerbations. An item is only counted as being present if the patient considers that it represents a break in normal functioning rather than a life-long trait (Goldberg 1972). Emphasis is therefore on changes in condition not on the absolute level of the problem. As a consequence, the GHQ may miss very long-standing disorders.

The GHQ-28

The 28-item GHQ was chosen for use in the studies presented in this thesis (Appendix 6). It is 'scaled' version of the 60-item GHQ, and was developed in 1979 by Goldberg on the basis of the results of principal components analysis. In comparison to other versions, where a single severity score is provided, it has the advantage of also providing a profile of scales scores on four dimensions - somatic symptoms, anxiety and insomnia, social dysfunction and severe depression. This allows identification of the areas in which problems are most concentrated, and which are most altered by an intervention (Goldberg and Hillier 1979).

2.7.4 Psychometric properties of the GHQ

There is considerable evidence to support the psychometric properties of the GHQ (Bowling 1995). The parent GHQ (60-item version) has been subjected to over 90 independent validations in more than 15 countries and has been translated into approximately 40 languages (Milne 1994).
1) Reliability

**Internal consistency**

High internal consistency has been demonstrated using the split-half technique, with high correlations demonstrated between the two halves for all versions (GHQ-60, \( r = 0.95 \); GHQ-30, \( r = 0.92 \); GHQ-28, \( r = 0.90 \); GHQ-12, \( r = 0.83 \)). Consistently high Cronbach’s alpha coefficients (\( \alpha = 0.8 - 0.93 \)) have also been shown for the 12-, 30-, and 60-item versions (Goldberg and Williams 1988).

**Interrater reliability**

Good interrater reliability for 12 interviews on the 30-item GHQ has been demonstrated with a disagreement of only 4% of symptom scores (Nott and Cutts 1982 cited in Goldberg and Williams 1988).

**Test-retest reliability**

It is methodologically difficult to measure this aspect of reliability since the GHQ is designed to assess a potentially highly variable quality (Goldberg and Williams 1988). As a consequence the definitive test-retest reliability study has not yet been undertaken. Results of a number of studies do indicate, however, that it is has adequate test-retest reliability. These studies include: a general population, after six months, when the stability of the patients condition was confirmed by repeating a standard psychiatric examination (\( r = 0.90 \), Goldberg 1978); a group of patients who judged their own condition as “having remained about the same” (\( r = 0.75 \), Goldberg 1978); immediate test and re-test scores in a neurological population (\( r = 0.89 \), De Paulo et al 1980); and a group of
stroke patients tested eight months apart \( r = 0.90 \), Robinson and Price 1982 as cited in Goldberg and Williams 1988).

2) Validity

This has been thoroughly tested by a number of studies, in a variety of different settings, using a range of techniques (McDowell and Newell 1987). Over 50 validity studies have been conducted on the GHQ (Bowling 1995), although it is noteworthy that few of these have been undertaken in physically disabled populations (Wade 1992).

Content validity

The initial pool of items for the GHQ was selected from existing instruments and clinical experience. The selection yielded 140 items which were then applied in a calibration exercise involving severely ill, mildly ill, and 'normal' people. On the basis of this work, 60 items, which had proven to be highly discriminatory between groups of patients, were selected for inclusion in the final version. Goldberg and Williams (1988) state that this method of test construction "guarantees" that the GHQ possesses content validity.

Criterion-related validity

Concurrent criterion validity

The psychiatric interview is the gold standard against which instruments for identifying psychiatric disorders are validated. Goldberg and Williams (1988) quote 22 studies which have shown the GHQ to correlate moderately to highly with the standardised psychiatric interview (range \( r = 0.53 - 0.83 \), median 0.7),
providing evidence to support its validity as a measure of emotional disturbance. The lower correlations were found when the GHQ was used in the community setting (Tarnopolsky et al 1979; Banks 1983), illustrating it may not be so useful in assessing severity within the ranges expected in community samples.

**Predictive criterion validity**

Although the GHQ was not developed as a predictive tool (Goldberg and Williams 1988), some studies have reported findings demonstrating predictive validity (predicting subsequent general practitioner consultations, and post-operative distress in breast cancer patients {cited in Wade 1992}). Two studies, however, have yielded negative results (Goldberg and Williams 1988).

**Construct validity**

**Convergent construct validity**

The GHQ-28 has shown to correlate highly with: total scores on the Clinical Interview Schedule (r = 0.76) and the Clinical Depression Rating (r = 0.73) in a primary care setting (Goldberg and Hillier 1979); the emotional reactions dimension of the Nottingham Health Profile (r =0.65, p<0.001, Jenkinson and Fitzpatrick 1990); the Clinical Interview Schedule in neurological inpatients (Lykouras et al 1996). Three of the four sub-scales on the GHQ-28 have also demonstrated moderate to strong correlations with relevant clinical ratings (Table 2.14, Goldberg and Hillier 1979).
Table 2.14 Associations between sub-scales GHQ-28 and clinical ratings

<table>
<thead>
<tr>
<th>Items from the Clinical Interview Schedule</th>
<th>GHQ Somatic</th>
<th>GHQ Anxiety / insomnia</th>
<th>GHQ Social dysfunction</th>
<th>GHQ Severe depression</th>
<th>GHQ Total score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Somatic symptoms</td>
<td>0.32</td>
<td>0.28</td>
<td>0.23</td>
<td>0.21</td>
<td>0.32</td>
</tr>
<tr>
<td>Anxiety and worry, 'anxious'</td>
<td>0.47</td>
<td>0.70</td>
<td>0.43</td>
<td>0.51</td>
<td>0.67</td>
</tr>
<tr>
<td>Despondency, 'depressed'</td>
<td>0.49</td>
<td>0.71</td>
<td>0.54</td>
<td>0.56</td>
<td>0.73</td>
</tr>
<tr>
<td>Psychiatrists severity</td>
<td>0.55</td>
<td>0.75</td>
<td>0.56</td>
<td>0.51</td>
<td>0.76</td>
</tr>
</tbody>
</table>

Associations measured by Spearman's correlation coefficient

With specific reference to the MS population, the GHQ-28 has been shown to correlate highly with the identification of emotionally distressed cases on clinical examination (misidentification rate = 8%), and the severity of depression as measured by the Present State Examination in outpatients ($r = 0.83$, $p < 0.001$, Rabins and Brooks 1981).

**Discriminant construct validity**

Items in the sub-scales of the GHQ-28 have proven to correlate less well with instruments measuring different dimensions (Table 2.14).

3) **Sensitivity and specificity**

Sensitivity is the proportion of true cases correctly identified by a test.

Specificity is the proportion of normals correctly identified at the same time.

(Tarnopolsky et al 1979). These features are important to consider if the GHQ is to be used as a valid screening instrument for detecting the presence of psychiatric morbidity in a population. Goldberg and Williams (1988) cited 12 studies which investigated this aspect of the validity of the GHQ-28 (Table 2.15).
Table 2.15  Validity studies of the GHQ-28

<table>
<thead>
<tr>
<th>Study</th>
<th>Setting</th>
<th>n</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Banks 1983</td>
<td>Unemployed adolescents</td>
<td>200</td>
<td>100</td>
<td>85</td>
</tr>
<tr>
<td>Bridges and Goldberg 1986</td>
<td>Neurological in-patients</td>
<td>70</td>
<td>80</td>
<td>81</td>
</tr>
<tr>
<td>Goldberg and Bridges 1987</td>
<td>Primary care attenders</td>
<td>283</td>
<td>87</td>
<td>75</td>
</tr>
<tr>
<td>Gold et al 1985</td>
<td>Accident and emergency department</td>
<td>27</td>
<td>56</td>
<td>88</td>
</tr>
<tr>
<td>Huxley et al 1987</td>
<td>Social work clients</td>
<td>140</td>
<td>84</td>
<td>76</td>
</tr>
<tr>
<td>Lindsay 1986</td>
<td>Amputees</td>
<td>35</td>
<td>100</td>
<td>86</td>
</tr>
<tr>
<td>Lobo 1986</td>
<td>Medical outpatients</td>
<td>100</td>
<td>84</td>
<td>82</td>
</tr>
<tr>
<td>Mann 1983</td>
<td>Schoolgirls</td>
<td>78</td>
<td>44</td>
<td>89</td>
</tr>
<tr>
<td>Medina-Mora et al 1983</td>
<td>Primary care attenders</td>
<td>352</td>
<td>73</td>
<td>74</td>
</tr>
<tr>
<td>Morris 1986</td>
<td>Gastroenterology patients</td>
<td>50</td>
<td>90</td>
<td>83</td>
</tr>
<tr>
<td>Rabins and Brooks 1981</td>
<td>MS patients</td>
<td>25</td>
<td>93</td>
<td>93</td>
</tr>
<tr>
<td>Sezlar and Mann 1987</td>
<td>Medical outpatients</td>
<td>22</td>
<td>100</td>
<td>75</td>
</tr>
</tbody>
</table>

(Reproduced from Goldberg and Williams 1988)

It is noteworthy that while the sensitivity is 100% in three of these studies (Banks 1983; Lindsay 1986; Selzer and Mann 1987), the number of true cases interviewed in each of these studies is very small (seven, seven and six respectively) and so their findings must be interpreted with caution.

Tarnopolsky et al (1979) found that sensitivity varied according to the prevalence of emotional disorders in the study population (ranging from 54% - 78%). As a consequence he suggested that the GHQ should be calibrated on a representative sample before adopting a "cut-off point" for a particular sample. For example, he suggested lowering the threshold in community settings where a lower proportion of cases are expected, and raising the threshold in populations with severe physical disability where the presence of somatic complaints may inflate the overall scores. A number of authors agree that
different cut-off points may be necessary for different population's (Bridges and Goldberg 1986; Minden and Schiffer 1991; Goldberg and Williams 1988; Lykouras et al 1996).

A threshold score of five points or more on the GHQ-28 and GHQ-30 has been used in MS studies to indicate probable emotional disturbance (e.g. Tarnopolsky et al 1980; De Paulo et al 1980; Rabins and Brooks 1981). The sensitivity and specificity rates for this cut-off point in the MS population have proven very high (92% for both sensitivity and specificity, Rabins and Brooks 1981). Results from a recent study of neurological inpatients (Lykouras et al 1996) confirmed high rates for sensitivity at this cut-off point (sensitivity = 87%) but found lower rates for specificity (specificity = 77%). The authors suggested that this higher rate of false positives may have been the result of an overinflation of scores due to somatic features.

4) Feasibility

The GHQ-30 and GHQ-28 has proven to be acceptable to patients, easy to score, and feasible for routine use in a neurological setting (De Paulo et al 1980; Rabins and Brooks 1981). A detailed operational and scoring manual, containing a comprehensive review of applications, and information regarding its psychometric evaluation, is available (Goldberg 1978; Goldberg and Williams 1988).

In summary, the GHQ-28 was chosen, in preference to other instruments, to measure the level of well-being and mood in the studies undertaken in this thesis. It's psychometric properties have been reasonably extensively
evaluated and it has proven clinically relevant and easy to administer. It has been used in studies to determine the prevalence and severity of emotional disturbance in MS, thereby allowing comparison of results.

2.8 Summary

The conceptual basis and terminology of the ICIDH underpins the theoretical and operational background to this thesis. A range of outcomes has been identified to determine the effectiveness of inpatient rehabilitation in patients with progressive MS (Table 2.16). The instruments used to measure these outcomes have been carefully selected according to their clinical and scientific properties (Table 2.16). They have been described in detail so that the results can be interpreted in the light of their strengths and weaknesses.

Table 2.16 Summary of outcomes assessed in this thesis

<table>
<thead>
<tr>
<th>Dimension assessed</th>
<th>Measurement Instrument</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impairment</td>
<td>Kurtzke's Expanded Disability Status Scale (EDSS) and Functional Systems (FS)</td>
</tr>
<tr>
<td>Disability</td>
<td>Functional Independence Measure (FIM)</td>
</tr>
<tr>
<td>Handicap</td>
<td>London Handicap Scale (LHS)</td>
</tr>
<tr>
<td>Health-related quality of life</td>
<td>36 item Short Form Health Survey Questionnaire (SF-36)</td>
</tr>
<tr>
<td>Well-being and mood</td>
<td>28-item General Health Questionnaire (GHQ)</td>
</tr>
</tbody>
</table>
Chapter 3

Outcome studies of comprehensive MS inpatient rehabilitation

Outcomes-based research is concerned with determining the results of health-care intervention and whether these results can be directly attributed to specific interventions (Frater and Costain 1992). As discussed earlier, this information is fundamental in enabling clinical decision making that is evidence-based (see Section 1.4).

Demonstrating the effectiveness of comprehensive rehabilitation is a complex task (Wade 1993). Different opinions exist as to how the outcome should be best evaluated (La Rocca et al 1994). One school of thought argues that the outcome of rehabilitation is determined by the cross-fertilisation and culmination of the multidisciplinary inputs involved in the "package" of care (Freeman et al 1996c). They consider that the evaluation of rehabilitation as a whole is therefore appropriate in reflecting the multidisciplinary rehabilitation process normally undertaken (La Rocca et al 1994), and that if the standard clinical package is not effective it is highly unlikely the component parts will be (Newham 1997). An opposing view is that no two centres practice in an identical manner, and therefore each component of the rehabilitation model (the 'parts') should be investigated separately. Both these approaches are valid, and contribute to the knowledge base of this subject (La Rocca and Kalb 1992; Pollock et al 1993). This thesis has focused on the evaluation of inpatient rehabilitation as a comprehensive "package of care".
3.1 Early Outcome Studies

Despite the large number of publications which describe the process of rehabilitation and provide anecdotal reports of its success (e.g. Slater 1980; Scheinberg et al 1981; Schapiro 1990; Mertin 1994; Thompson et al 1994), there have been relatively few objective studies of this subject (La Rocca et al 1994). A review of the literature revealed seven studies, published in peer reviewed journals, which investigated the outcome of comprehensive MS inpatient rehabilitation (Feigenson et al 1981; Greenspun et al 1987; Reding et al 1987; Carey et al 1988; Francabandera et al 1988; Kidd et al 1995a; Aisen et al 1996). Of these, four were undertaken at the same centre (Feigenson et al 1981; Reding et al 1987; Francabandera et al 1988; Aisen et al 1996). Recently the results of a further three studies have been published in abstract (Cendrowski et al 1995; Vaney et al 1996) and letter format (Kidd and Thompson 1997, in press). These studies (Table 3.1) have asked a variety of questions in their evaluation of comprehensive inpatient rehabilitation.

i) Is comprehensive MS inpatient rehabilitation effective?

Feigenson and colleagues (1981) were among the first to publish evidence about the efficacy of comprehensive inpatient rehabilitation in MS. They used a prospective, single group, pre- and post treatment design. The sample consisted of twenty patients, with long-standing disability, all of whom had failed to respond to regular intensive multidisciplinary outpatient treatment. Nineteen of the 20 patients were female with a mean age of 44.8 years (range 22-62), and disease duration of 12.9 years (range 1-25 years). Following an average length of stay of 4.0 days, statistically significant improvements were
noted in seven different functional areas: balance, self care activities, bed mobility, transfers, wheelchair management, homemaking skills and real-life activities. In addition the amount of help required to maintain patients at home was "substantially decreased" for the entire year following rehabilitation. No significant change in any impairments was demonstrated, suggesting that spontaneous natural recovery (or deterioration) in neurological dysfunction had not occurred during the study period. Cost-benefit analysis was also undertaken. Following an initial assessment, crude estimates of the cost of care prior to rehabilitation was made by the treating team. The cost of rehabilitation was abstracted from invoices. One year following rehabilitation patients were surveyed by telephone to estimate their current level of home assistance. Comparison of these figures showed a reduction in the annual cost of care following rehabilitation from $26,153 to $10,583. While acknowledging the crude method of cost estimation, the authors appeared confident that inpatient rehabilitation was cost-effective.

Greenspun and colleagues (1987) undertook a retrospective, single group, pre-and post treatment study to investigate the effectiveness of inpatient rehabilitation on reducing disability. The Computerised Rehabilitation and Data System (CRDS) was used to evaluate disability on admission, discharge and at three months follow-up. Data was gathered on twenty eight patients (75% female, mean age 42 years, mean disease duration 12.2 years), who participated in 33 episodes of inpatient rehabilitation over a four year period. The average length of stay was 28 days (range 5 - 57). Neurological status was not monitored throughout the study period. Results were reported by describing
the change in the percentage of patients improving their level of independence in mobility and self care activities. Between admission and discharge, independence in ambulation improved from 18% to 76% of patients, and in stair climbing from 9% to 64% of patients. Independence in self care activities improved from: 54% to 73% in dressing; 39% to 70% in bathing; 70% to 85% in toileting; and 91% to 94% in eating. At follow-up three months later, the majority of these improvements had been maintained.

The largest published study was undertaken by Carey and colleagues (1988). This retrospective, single group, pre- and post treatment study gathered data from 22 rehabilitation units. Its aim was to determine who made the most functional gains in inpatient rehabilitation. This study was not specific for MS, but studied eleven different diagnostic groups. Assessments of disability were made on admission and discharge. Comparisons were made between MS patients and other conditions. On average, the MS patients were less disabled on admission as measured by the Revised Level of Rehabilitation Scale (LORS-II), and had a shorter length of stay (22 days versus means ranging from 28 to 46 days). Patients with MS were younger than other diagnostic groups (mean 46 years versus means ranging from 49 - 70 years) apart from head injured patients (mean 37 years). Neurological status was not monitored. Results showed that MS patients improved their level of performance on the LORS-II. These gains however were smaller than those of all other diagnostic groups.
The first study to distinguish patients by disease group, and to identify those recovering from a relapse or taking steroids, was undertaken by Kidd and colleagues (1995a). They used a prospective, single group, pre- and post treatment design to study 79 patients (49 females; mean age 48.8 years, range 17-61; mean disease duration 12.1 years, range 1-37). Assessments were made on admission and discharge using measures of impairment (Kurtzke's DSS), disability (Barthel Index) and handicap (ESS). In terms of disease pattern, 15 patients were relapsing/remitting, 57 secondary progressive, and seven primary progressive. The severity of disease was reflected by the relatively high scores on the DSS (median 7.0, range 4.0-9.0). The length of stay was short (mean 15 days, range 1-59). Following rehabilitation, statistically significant improvements were demonstrated by 65% of patients in disability, and by 44% in handicap. Improvement was most marked in those in whom a reduction in impairment had occurred, but was also seen in 63% of those who had not changed neurologically.

Do the benefits gained carry over in the longer term?

Each of the previous studies looked at the short term benefits of inpatient rehabilitation. Two studies attempted to determine whether these benefits gained were carried over following discharge into the community. Aisen and colleagues (1996) designed a retrospective, single group, pre- and post treatment study to address the question of long term carryover. The authors reviewed the notes of 37 consecutive MS patients who were admitted to rehabilitation following a functional decline. Assessments of impairments (EDSS and FS scales) and disability (FIM) were reviewed on admission and
discharge. In addition patients were followed up by interview (at nine months) and by telephone (once or twice somewhere between 6-36 months). Ninety percent of subjects were female, with an average age of 47 years (range 24-68), time since diagnosis of 11.8 years (range 0.1 - 32), and EDSS score of 7.5. In terms of disease pattern, six patients were described as relapsing remitting, five relapsing progressive, and 26 chronic progressive. The average length of stay was 32 days (range 12-77). Statistically significant improvements between admission and discharge were noted in all outcomes irrespective of disease pattern. Pyramidal and cerebellar functions improved particularly well. It is difficult to determine how long these benefits were maintained following discharge since the assessments were not performed at standardised points over the 6-36 month follow-up period.

In a recently to be published letter, Kidd and Thompson (1997, in press) described a prospective study evaluating the outcome at three months following discharge. A single group, pre- and post treatment design was utilised. Forty seven patients were assessed on admission and discharge (75% female, mean age 40, median disease duration 13 years). Of these, forty four were followed up at three months. Assessments, at each time point, included impairment (EDSS), disability (FIM) and handicap (ESS). At discharge improvements were demonstrated by 17% of patients in impairment, 87% in disability and 47% in handicap. At three months, eleven patients had suffered new neurological symptoms, sufficient to cause a deterioration in the EDSS score in two patients. Gains in disability had been partly maintained in 86% of patients. Handicap further improved following discharge with an additional 30%
of patients (total patients improved 77%) improving their handicap score at three months. The authors suggested that the continued improvements in handicap were reflective of the work planned during the rehabilitation programme which had been subsequently carried out in the community.

*How does inpatient rehabilitation compare to other forms of intervention?*

This question has been addressed by two studies (Reding et al 1987; Francabandera et al 1988). Reding and colleagues (1988) compared the outcome and cost of inpatient rehabilitation to acute hospital care using retrospective case matched analysis. Data from 20 pairs of patients, matched for sex and severity of MS, was reviewed. The level of disability (Incapacity Status Scale), rate of re-hospitalisation, and need for home help assistance was determined by telephone review at 16 months follow-up. Costs of care were crudely estimated. Changes in neurological status were not monitored. At 16 month follow-up, no difference was demonstrated between the two groups in either their functional status or rate of re-hospitalisation. Although the cost of inpatient rehabilitation was less per day than the acute hospital setting ($US364 versus $US625), the overall cost of admission was greater because of the longer length of stay (35 days versus 14 days). In reviewing this study, La Rocca and Kalb (1992) suggested that if the average length of rehabilitation stay could be shortened, savings could be made without compromising patient care.
Francabandera and colleagues (1988) compared inpatient to outpatient intervention. Following random assignment to either group, 84 patients were assessed on the level of disability (Incapacity Status Scale) prior to admission and then at three monthly interviews for two years after discharge. The hours of home assistance required by the patient were recorded. On admission the two groups were comparable in terms of age, sex, EDSS scores, and hours of home assistance required. The initial level of disability between the groups was however statistically different - the inpatient group was more disabled than the outpatient group. Analysis of covariance was used to control for this. Comparison of discharge scores showed significant improvements in disability in the inpatient group, in contrast to the outpatient group where no improvements occurred. Preliminary results from the three month follow-up demonstrated a statistically significant difference between the two groups. The inpatient group had improved in terms of disability whereas those in the outpatient group had deteriorated slightly. No impact on home care was shown. Neurological status was not monitored. A brief summary of results from the 12 month follow-up was later reported by La Rocca and Kalb (1992). By twelve months both groups had reverted to pre-treatment levels of disability, with no statistically significant differences between them. They concluded that the initial advantage gained by inpatient rehabilitation appeared to be short-lived, and suggested that a periodic course of rehabilitation may therefore be necessary.
<table>
<thead>
<tr>
<th>Reference</th>
<th>Trial Method</th>
<th>Sample (n)</th>
<th>Main outcomes / instruments</th>
<th>Timing of assessments</th>
<th>Results</th>
<th>Authors conclusions</th>
</tr>
</thead>
</table>
| Feigenson et al 1981 | Prospective, single group, pre- and post study design. Data gathered from multidisciplinary team assessment. | 20         | Impairment, disability and handicap: MS Functional Profile (a modified version of BUSTOP). Costs of intervention abstracted from invoices. | Admission and discharge. Costs measured at discharge and 12 months (by telephone review). | Average length of stay: 40 days (sd 11)  
Statistically significant improvements were noted in: balance (p< 0.001); self-care activities (p<0.000); bladder control (p<0.006); bed mobility (p<0.000); wheelchair transfers (p<0.000); ambulatory transfers (p<0.041); homemaking (p<0.000) and real-life activities (p<0.001)  
No significant changes were noted in primary neurological impairments (sensation, spasticity, cognition, communication, incoordination).  
Average estimated cost of rehabilitation per person = $14,175. At 12 months review the average home services costs reduced from $26,153 (pre admission) to $10,583 (post rehabilitation). | Inpatient rehabilitation was associated with a significant reduction in disability, despite fixed neurologic deficits. It was suggested that these gains could not be realised on an outpatient basis.  
Intensive rehabilitation did not enable nonambulatory patients to walk. It was hypothesised that inpatient rehabilitation may diminish the costs of home services by 2/3rds during the first year after treatment, with similar savings in subsequent years. |
| Greenspun et al 1987 | Retrospective, single group, pre- and post study design. Data gathered over a 4 year period. | 28         | Disability: CRDS.  
|                |                               |            | Admission, discharge, and 3 month review (by telephone if necessary).                       |                                                                                  | Average length of stay: 28 days (range 5-57). Changes in function were noted in independence in: ambulation (8% on admission, 76% by discharge); stair climbing (9% on admission, 64% by discharge); transfers (58% on admission, 82% by discharge); dressing (54% on admission, 73% by discharge); bathing (40% on admission, 70% by discharge); toileting (70% on admission, 85% by discharge) and eating (91% on admission, 94% by discharge)  
 Functional gains made in all areas were maintained at 3 month review | Inpatient rehabilitation was associated with a reduction in disability. The gains made were well maintained at 3 month review.  
Length of stay and functional ability at admission were important factors in patient progress. In general, patients who had a longer length of stay and were initially more dependent made greater relative gains in function. |
| Reding et al 1987 | Retrospective study, using case matched analysis. Data gathered from chart and telephone review. | 20 pairs   | Disability: ISS Hospital re-admission rate. Cost of intervention: abstracted from invoices The need for home help assistance. | Review at 16 months (by telephone).                                            | Average stay at rehabilitation unit (approximated): rehabilitation group 35 days acute hospital care group 14 days. Estimated cost per diem: acute hospital care = $625 rehabilitation care = $364.  
At 16 month review no significant differences in disability status (p=0.43) or hospital re-admission rate were noted between the 2 groups. | It was hypothesised that an estimated 50% of people admitted for acute hospital care could have been equally well managed in a less expensive rehabilitation setting. It was suggested that a significant reduction in hospital cost might be achieved if patients were admitted to rehabilitation, for a carefully controlled length of admission (<2-3 weeks), rather than acute hospital. This reduction in cost is dependent on the length of stay |
<table>
<thead>
<tr>
<th>Reference</th>
<th>Trial method</th>
<th>Sample (n)</th>
<th>Main outcomes / instruments</th>
<th>Timing of assessments</th>
<th>Results</th>
<th>Authors conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carey et al 1988</td>
<td>Retrospective multi-centre study assessing a range of conditions. Single group, pre- and post study design. Data gathered over a 2 year period.</td>
<td>6,194 inpatients, of whom 196 had MS.</td>
<td>Disability :LORS-II</td>
<td>Admission and discharge</td>
<td>For MS patient group: Average length of stay: 22 days (sd 17) Changes were noted in functional activities of daily living (mean gain 16 points) and mobility (mean gain 20 points)</td>
<td>Inpatient rehabilitation was associated with a reduction in disability. Compared to other diagnostic groups, patients with MS: - had the shortest average length of stay, - made the least gain in ADL Length of stay, age, and functional ability at admission were important factors in patient progress. In general, MS patients who were older, were initially in the middle ranges of dependence for function, and had a longer length of stay made greater relative gains in function.</td>
</tr>
<tr>
<td>Francabandera et al 1988</td>
<td>Prospective study. Stratified randomised assignment to either inpatient or outpatient rehabilitation. Data gathered by either telephone or clinic interview.</td>
<td>84 patients with an EDSS of between 6.0 - 9.0</td>
<td>Disability: ISS Need for home assistance: (hours)</td>
<td>Admission, and at 3 monthly intervals for 2 years. N.B. Only the results of the three month follow-up are reported.</td>
<td>Preliminary analysis (results of 3 month follow-up): A slight but statistically significant difference was noted between the groups at 3 month review, with improvement in ISS scores in the inpatient group but deterioration in the outpatient group. No significant difference was noted in the need for home assistance between the groups</td>
<td>Inpatient rehabilitation was associated with slight improvements in functional ability, compared to the outpatient group where no improvements were measured. Longer term follow-up is needed to determine whether these improvements will be sustained.</td>
</tr>
<tr>
<td>Kidd et al 1995a</td>
<td>Prospective study, single group, pre- and post study design. Data gathered by multidisciplinary team assessment.</td>
<td>79</td>
<td>Impairment: DSS Disability: Barthel Index Handicap: ESS</td>
<td>Admission and discharge</td>
<td>Average length of stay :15 days (s.d.11) Statistically significant improvements were noted in: disability in 65% of patients; handicap in 44% of patients. Impairment remained unchanged in 63% of patients.</td>
<td>Inpatient rehabilitation was associated with a reduction in disability. Category of disease pattern and functional ability at admission were important factors in patient progress. In general, patients who demonstrated a reduction in impairment, and those who were initially in the middle ranges of dependence for function made greater relative gains in disability</td>
</tr>
<tr>
<td>Cendrowski et al 1996 Abstract</td>
<td>Multicentre study (? design, ? method of data collection). Comparisons made between multidisciplinary rehabilitation, and physical exercises.</td>
<td>588</td>
<td>Disability: EDSS</td>
<td>Not stated</td>
<td>60% (range 46-82%) had an immediate reduction in physical disability. Repeated courses of rehabilitation over 4-4.5 years was more effective in maintaining improvement (37% and 55% patients improved), compared to a single episode (26% improved). Multidisciplinary intervention was more effective (82% improved) than physical exercises (60% improved).</td>
<td>Recurrent inpatient rehabilitation temporarily improves the quality of life.</td>
</tr>
</tbody>
</table>
### Table 3.1 continued Outcome studies of comprehensive MS inpatient rehabilitation.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Trial method</th>
<th>Sample (n)</th>
<th>Main outcomes / instruments</th>
<th>Timing of assessments</th>
<th>Results</th>
<th>Authors conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aisen et al 1996</td>
<td>Retrospective, single group, pre- and post study design. Data gathered from chart and telephone review.</td>
<td>37</td>
<td>Impairment: EDSS Disability: FIM</td>
<td>Admission, discharge and telephone follow up (from between 6-36 months post discharge)</td>
<td>Average length of stay: 32 days (s.d. 8.73, range 12 - 77 days)</td>
<td>Significant improvements were noted in: FIM (p=0.0001); FS (p=0.0001); and EDSS scores (p = 0.0001) for all groups combined between admission and discharge. In terms of disability, significant improvements were noted in 3 FIM subscales: self-care (p=0.0001); sphincter control (p=0.0222); locomotion (p=0.0001). In terms of impairment, significant improvements were noted in 3 FS scales: pyramidal (p=0.001); cerebellar (p=0.0033); sphincter control (p=0.048). Gains were partly maintained between discharge and 6-12 month follow-up (n=17). Deterioration was evident in FIM scores at 24-36 months review (n=12).</td>
</tr>
<tr>
<td>Vaney et al 1996</td>
<td>Prospective, single group, pre- and post study design.</td>
<td>444</td>
<td>Disability: FIM</td>
<td>Admission and discharge</td>
<td>Median length of stay: 27 days</td>
<td>Overall, FIM scores showed that 25% of patients improved, 4% deteriorated, and 71% remained unchanged. Of those improving, 25% changed their score by 1 FIM point. According to EDSS score improvements in FIM was most marked in patients with an EDSS of 8.0-8.5 (37% improved) and least in EDSS scores of 3.0 - 6.0 (15% improved) and 9.0-9.5 (12% improved).</td>
</tr>
<tr>
<td>Kidd and Thompson 1997</td>
<td>Prospective, single group, pre- and post study design</td>
<td>47</td>
<td>Impairment: EDSS Disability: FIM Handicap: ESS</td>
<td>Admission, discharge and 3 month follow-up</td>
<td>At discharge (n=47), improvements were demonstrated by 17% of patients in impairment; 87% in disability; 47% in handicap. At 3 months (n=44) gains in disability had been partly maintained in 86% of patients. A further 30% of patients had improved their handicap score (total 77%).</td>
<td>Inpatient rehabilitation was associated with a reduction in disability and handicap.</td>
</tr>
</tbody>
</table>

BI = Barthel Index; BUSTOP= Burke Stroke Time-oriented Profile; DSS= Disability Status Scale; EDSS = Expanded Disability Status Scale; ESS = Environmental Status Scale; FIM = Functional Independence Measure; LORS-II = Revised Level of Rehabilitation Scale; FS= Functional Systems; ISS = Incapacity Status Scale.


*What aspects of rehabilitation are effective?*

Little knowledge has been gained from these studies about what problems are most amenable to change and what elements of the rehabilitation programme are effective in achieving these benefits. These questions are best addressed by evaluating specific aspects of process and outcome (the 'parts') rather than the comprehensive 'package' of care. Recently, two randomised controlled trials have been published which have investigated specific components of the rehabilitation package (Petajan et al 1996; Fuller et al 1996).

Petajan and colleagues (1996) assessed the effect of a 15 week aerobic training programme on aspects of fitness, daily activities, mood, fatigue and disease status in 54 MS patients who were randomly assigned to either the exercise or non-exercise group. Compared with baseline the control group showed no significant changes over time for any variable. In contrast, the exercise group demonstrated statistically significant improvements in the short term (up until 10 weeks) in maximal aerobic capacity, isometric strength, body composition, blood lipid levels, depression and anger and fatigue. By 15 weeks, however, many of the psychological benefits were no longer significant. While acknowledging that some of the benefits may have resulted from the increased attention and social interaction, the authors concluded that exercise training improved fitness and had a positive impact on quality of life.

The effectiveness of a single inpatient admission for physiotherapy on improving mobility and related activities of daily living was investigated by Fuller and colleagues (1996). Forty five MS patients, who had a recent history
of deterioration in gait or transfer ability, were randomised to either “early” (immediate physiotherapy) or “late” treatment (physiotherapy delayed by nine weeks). Patients in the early group participated in physiotherapy for an average of 38 minutes per working day for 13.5 days. Assessments of mobility, activities of daily living, subjective visual analogue scores (patient and carer) and rating of randomised video clips of mobility were undertaken at baseline and nine weeks. No statistically significant differences in change scores were detected in either group for disability, mobility, or video clip ratings. A significant reduction in mobility-related distress, as measured by the patient visual analogue scale, was noted in the treatment group. The authors suggested that improved patient selection or specific goal-directed intervention may improve the efficacy of the physiotherapy programme.

What are the conclusions of the studies evaluating comprehensive rehabilitation?

These uncontrolled studies do not provide convincing evidence of the effectiveness of MS inpatient rehabilitation. Furthermore, it is difficult to make meaningful comparisons between the studies due to the different sample populations and methodologies utilised. Few have described the sample population in terms of disease severity, many have failed to assess ongoing neurological status, and a wide range of outcome measures have been used. Unlike drug trials where patient selection is an important feature of trial design, only one study identified specific entry and exclusion criteria. These issues are discussed in detail in Section 3.2.
It is notable, however, that every study demonstrated a reduction in the overall level of disability immediately following rehabilitation. Three also showed that it was associated with reduced levels of handicap (Feigenson et al 1981; Kidd et al 1995a; Kidd and Thompson 1997, in press). Some differences in outcomes are evident. For example, whereas Feigenson and colleagues (1981) reported that non-ambulators remained non-ambulators, Greenspun and colleagues (1987) reported substantial gains in ambulation. Marked differences in the mean length of stay are also noted, varying from an average of 15 days (Kidd et al 1995a) to 40 days (Feigenson et al 1981).

In terms of carryover, the limited evidence available suggests that the benefits gained are partly maintained in the longer term (Aisen et al 1996; Kidd and Thompson 1997, in press), although a return to initial levels of performance within twelve months was shown (Francabandera 1988). Conclusions about cost benefit cannot be made. In each of the two studies analysing this aspect (Feigenson et al 1981; Reding et al 1987) costs were determined retrospectively, and were crudely estimated. Their findings should be interpreted with caution.

3.2 Current methodological dilemmas

A number of methodological problems are encountered when designing clinical studies both in the fields of rehabilitation (Tallis 1989; Pollock et al 1993; Wade 1993) and MS (Ellison et al 1994; Whitaker et al 1995; Rudick et al 1996). Scientifically sound research which quantifies effectiveness, while remaining clinically relevant, is therefore particularly challenging.
Some of these methodological difficulties include:

i) Measurement of change

The chronic, progressive, variable and unpredictable nature of the disease course means that there is no stable point from which to make reference baseline measurements (Figure 3.1). Even the most recent predictive methods (Runmarker et al 1994; Weinshenker et al 1996) have proven unreliable for determining the likely pattern or rate of deterioration (or recovery) in individual patients. Further problems arise because patients are admitted for rehabilitation at different stages of the disease process and at different starting levels of disability (Langdon and Thompson 1995). These issues render the interpretation of repeated assessments a formidable task.

Figure 3.1 Pattern of disease activity in multiple sclerosis in most patients

(Reproduced with permission from the artist and author AJ Thompson, 1996b).
This has led authors to recommend obtaining data from control groups (e.g. Ellison et al 1994; Whitaker et al 1995), and measuring the level of neurological status at each re-assessment (Rintala 1987). To date, however, no controlled studies have been undertaken of comprehensive MS inpatient rehabilitation; many have failed to assess the patients ongoing neurological status (e.g. Greenspun et al 1987; Reding et al 1987; Carey et al 1988; Francabandera et al 1988; Cendrowski et al 1996); and all have included patients in both the relapsing remitting and progressive phase of the disease. Consequently, it is difficult to determine from these studies whether improvement was a consequence of the inpatient rehabilitation intervention, spontaneous neurological recovery from relapse, or the natural history of the disease.

**ii) The nature of comprehensive inpatient rehabilitation**

A clearly described and standardised intervention is required in order to conclusively attribute the outcome to a specific intervention (Wilkin et al 1992). Comprehensive inpatient rehabilitation, however, is an all embracing concept with broad ranging goals and varying interventions (refer to Section 1.3.1). Recognition of the diverse and ever-changing problems of people with MS means that programmes are tailored to the needs of individuals, adapting treatment to changes as they occur. Precisely defined standardised interventions are hence not only difficult, but often inappropriate, in a clinical context.
iii) Choice of appropriate outcomes

Choosing and measuring outcomes which accurately reflect both the problems of MS (Hobart and Thompson 1997) and the broad aims of the rehabilitation programme (Jeffrey 1993) is difficult. As discussed earlier, the aims of rehabilitation are many and varied, and are concerned with much more than simply the patient's capacity for self care. The majority of studies, however, have focused on the measurement of impairment and/or disability (Greenspun et al 1987; Carey et al 1988; Aisen et al 1996; Cendrowski et al 1996; Vaney et al 1996). This approach risks missing valuable benefits - particularly in relation to the patient perspective (Hobart et al 1996a).

iv) Use of appropriate measurement instruments

The use of clinically relevant, psychometrically evaluated instruments, is essential to confidently interpret study results (refer to Section 2.2). Currently however there is a paucity of comprehensively evaluated instruments which have been found suitable to measure the primary aims of the rehabilitation programme (Wade 1992; Jeffrey 1993). As a consequence, researchers have been forced to use instruments, available at the time the study is initiated, which have not been comprehensively evaluated (Rudick et al 1996) and are imperfect (Johnston et al 1992; MS Forum Workshop 1996).

In addition, there is a lack of agreement as to which instruments best assess the results of rehabilitation. As a consequence, a wide variety of scales have been used to assess the same outcome. For example, of the ten studies evaluating MS rehabilitation, six different instruments were used to measure
disability. This further contributes to the difficulty in interpreting, comparing and contrasting results.

Knowledge of the reliability of the outcomes measured in studies is important (Wilkin et al 1992). A number of the studies gathered outcomes data retrospectively and consequently the reliability of the results was not determined (Greenspun et al 1987; Reding et al 1987; Carey et al 1988; Aisen et al 1996). The prospective studies have tended to use multiple raters, but have not reported information about the level of inter-rater reliability.

v) Attributing the outcome to the intervention

The large number of variables involved in MS rehabilitation makes attribution of the outcome to the intervention extremely complex (Theriot and Brar 1993). Controlling within a trial for these confounding factors is often difficult, and at times impossible (Pollock et al 1993). For example, the accomplishment of many of the goals of rehabilitation depend on factors outside the control of the rehabilitation programme. These include: family support, access to community services, provision of aids and equipment, and economic considerations. If these influences are working against patients, even the best conducted programme is vulnerable to not meeting its goals (Fuhrer 1987). The converse can also be true. This means attribution to the intervention is not always clear.

Incorporation of a placebo control group into the study design is recommended to resolve some of these problems (Whitaker et al 1995). A number of factors may have prevented this from being included in previous trials. Firstly, the
placebo for rehabilitation has not yet been defined; secondly, both practical and economic constraints considerations make this difficult to achieve.

A further complexity, with regards to attribution, is that the consequences of the service may take some time to become apparent. Knowledge of the carryover of changes once the patient is discharged home is crucial. Disregarding this can lead to a misinterpretation of the effectiveness of interventions, either suggesting the programme had not been effective when it had, or that it was effective when review in the longer term demonstrated otherwise (Fuhrer 1987). No guidelines currently exist as to the appropriate times over which interventions should be evaluated.

Relatively few studies in MS have incorporated follow-up reviews in their design. While Reding and colleagues (1987) investigated the level of disability and amount of home care required at 16 months post discharge, they did not report any initial baseline measurements. It is therefore impossible to determine whether changes had occurred. Feigenson (1981) and Francabandera (1988) both included a twelve month review, however Feigenson limited his review to a re-assessment of costs, and only the three month results of Franacabandera’s study has been published in detail. Two studies have assessed the changes carried over at three months (Greenspun et al 1987; Kidd and Thompson 1997, in press).
vi) Heterogeneity of the study population

Homogeneity of sample populations is recommended when designing outcome studies. This is difficult to achieve in MS, where a diversity of clinical manifestations is observed both in terms of disease course and clinical presentation (Rudick et al 1996). Unlike drug trials, where patient selection is an important feature of the study design (MS Forum Workshop 1996), no agreed selection criteria exist for inpatient rehabilitation. To date, all of the earlier studies have included patients in both the relapsing remitting and progressive phase of the disease, and only one identified specific entry and exclusion criteria (Francabandera et al 1988).

The heterogeneity of MS means that definitive studies are likely to require large numbers of patients (Hobart and Thompson 1997). This has proven difficult. To date the sample size of most studies has been small, with only one trial incorporating the multicentre methodology necessary to recruit large numbers of patients (Cendrowski et al 1996).

These factors illustrate why research in the gold standard format of double-blind randomised placebo-controlled studies is extremely difficult to achieve in neurorehabilitation. It partly explains why relatively few studies have been attempted in this subject.

3.3 Rationale for the studies undertaken in this thesis

There have been many calls to improve the quality of measurement and evaluation procedures of MS inpatient rehabilitation (La Rocca et al 1994;
Johnson 1996; Thompson 1996c). As discussed, a number of methodological difficulties have been encountered in the earlier studies, some of which remain unresolved. Current methodological dilemmas include:

- inadequate comparison (control) groups
- retrospective research design
- heterogeneous study samples
- failure to monitor levels of impairment and disease severity
- lack of blinding
- small sample sizes
- inadequate follow-up assessments to determine the extent of carryover
- decreased generalisability due to single centre studies
- lack of evaluation from a patient perspective
- use of a wide range of measurement instruments, many of which have not been comprehensively evaluated
- inadequate description of the methodology, intensity and duration of rehabilitation intervention.

Nevertheless the earlier studies have provided a foundation from which future research may develop. The intention of this thesis is to contribute to this area of research by addressing some of these unresolved methodological problems. The basis of the thesis is the design of two, closely linked studies. The first study has a randomised controlled design which investigates the effectiveness of MS inpatient rehabilitation in the short term. The second study utilises a longitudinal design to observe the carryover of changes in the longer term.
4.1 Aim of the Study

The aim of this study is to determine whether multidisciplinary inpatient rehabilitation results in improved levels of impairment, disability, handicap, health-related quality of life, or emotional well-being in patients with progressive MS.

4.2 Null Hypothesis

Patients with progressive MS participating in intensive multidisciplinary inpatient rehabilitation will have the same clinical outcome as control patients in terms of impairment, disability, handicap, health-related quality of life and emotional well-being.

4.3 Study Design

Prospective stratified randomised controlled study.

4.4 Subjects and Methods

Ethical approval was obtained from the Ethics Committee of the National Hospital for Neurology and Neurosurgery in April 1994.

4.4.1 Sample size

Seventy patients were recruited into this study. The sample size was estimated from the unit’s audit data which records the patient’s disability and handicap.
scores on admission and discharge from rehabilitation. The calculation was based on a two-sided significance test at 80% power and 0.05 significance.

Since no controlled trials have been undertaken in this area, and a variety of outcome measures have been used in the few uncontrolled trials, it was impossible to use the results from earlier studies to undertake precise power calculations.

4.4.2 Recruitment

Subjects were recruited from the weekly multidisciplinary assessment clinic at the National Hospital for Neurology and Neurosurgery (NHNN), Queen Square. Assessment is routinely performed by a multidisciplinary team comprising a consultant neurologist, superintendent III physiotherapist, head III occupational therapist, and a clinical nurse specialist, with the addition of a head speech therapist and clinical psychologist when required. Patients are accepted for inpatient rehabilitation if they are medically stable and considered "likely to benefit" from rehabilitation involving two or more disciplines. The anticipated benefits may be either in terms of functional improvement or the organisation of an appropriate package of community support to maintain the patient in their own home. (Johnson and Thompson 1996).

i) Inclusion criteria

To exclude as many confounding variables as possible, patients were only entered into the study if they had a diagnosis of clinically or laboratory supported definite MS (Poser et al 1983), were in the progressive phase of the disease as established by the neurologist (Lublin and Reingold 1996), and
were considered appropriate for inpatient rehabilitation (Freeman et al 1996a; Johnson and Thompson 1996).

**ii) Exclusion criteria**

Patients were excluded if they were within one month of relapse, were within one month of receiving steroids, required urgent admission on clinical grounds, had other diseases which may have interfered with outcome, or were cognitively impaired such that they were unable to give informed consent.

4.4.3 Procedures

All patients who met the entry criteria were informed of the study by both written and verbal explanation (Appendix 7). If agreeable to participation, written consent was obtained (The NHNN's "Patient Consent Form for Research on Human Volunteers").

4.4.4 Stratification and randomisation

To maximise comparability between the treatment and control groups a stratification procedure, in relation to disease severity, was undertaken prior to randomisation. This stratification process involved categorising patients into three levels of disease severity, to ensure equal representation of each level in the two groups. All patients recruited were assessed by a single neurologist who scored them using Kurtzke's EDSS. They were then categorised into one of three levels of disease severity according to the EDSS score: mild (EDSS = 0.0 - 4.5), moderate (EDSS = 5.0 - 6.5) or severe (EDSS = 7.0 - 9.5) (BSMR Working Party Report on MS 1993). Subsequently patients within each band of
disease severity were randomly allocated, by means of a random numbers table (from Armitage and Berry 1987), to either immediate treatment (treatment group) or a waiting period of six weeks during which time no rehabilitation intervention was provided and no other interventions were arranged (control group).

Patients randomised to the treatment group were assessed within 24 hours of admission to the neurorehabilitation unit (NRU) and then again six weeks later, by which time all had completed the rehabilitation programme and returned home. Those randomised to the control group were assessed immediately in the clinic and then again six weeks later, having had no rehabilitation intervention (Figure 4.1).

**Figure 4.1 Outline of study with sequence of assessment events.**

<table>
<thead>
<tr>
<th>CLINIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assessment 1</td>
</tr>
<tr>
<td>on admission for rehabilitation</td>
</tr>
<tr>
<td>Rehabilitation</td>
</tr>
<tr>
<td>6 WEEKS</td>
</tr>
<tr>
<td>No Rehabilitation</td>
</tr>
<tr>
<td>Assessment 2</td>
</tr>
<tr>
<td>6 weeks following admission</td>
</tr>
<tr>
<td>Assessment 1</td>
</tr>
<tr>
<td>at the outpatient clinic</td>
</tr>
<tr>
<td>Assessment 2</td>
</tr>
<tr>
<td>6 weeks later, on admission to rehabilitation</td>
</tr>
</tbody>
</table>
4.4.5 Test battery and methods of administration

i) Demographic and diagnostic information

The following information was collected by patient interview and verified by review of the medical notes:

a. Routine demographic and diagnostic data (Appendix 8).

b. Number of relapses in the six months preceding and during the study period. A relapse was defined as the occurrence of a symptom or symptoms of neurological dysfunction, with or without objective confirmation, lasting more than 24 hours (Poser et al 1983).

c. Pharmacological details.

Drug regimes, and alterations to them during the six week study period, were recorded. Drugs were assigned to one of 10 categories (Table 4.1), a system devised by the Consultant Neurologist, Dr Alan Thompson, to enable analysis of relevant changes.

<table>
<thead>
<tr>
<th>Drug Category</th>
<th>Immunsuppressants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Steroids</td>
<td>Anti-depressants</td>
</tr>
<tr>
<td>Anti-spasticity agents</td>
<td>Fatigue management</td>
</tr>
<tr>
<td>Anti-tremor agents</td>
<td>Bowel management</td>
</tr>
<tr>
<td>Bladder management</td>
<td>Other</td>
</tr>
<tr>
<td>Analgesics/Sensory disturbance</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.1 Drug categorisation system

ii) Measures

All patients were assessed by each measure at zero weeks and six weeks. At each assessment the measures were administered in a randomised order (from
random numbers table, Armitage and Berry 1987), to prevent any systematic bias resulting from practice effects or fatigue (Daly et al 1991). To minimise bias, the assessors did not have further access to any of the initial scores before the second assessment. Similarly, patients were not permitted to see their responses from previous occasions.

The following measurement instruments were used:

   Assessment of neurological status (impairment) was determined by Kurtzke's FS and EDSS. The former comprises eight scales, each measuring a specific function within the central nervous system. Scores for each scale range from zero to six where zero is normal and six is maximal impairment. The EDSS is a single index of severity of MS ranging from zero (normal) to 10 (death). While the lower section of the scale addresses impairment, grades four to ten are strongly dependent on one aspect of disability, locomotion. Scores on both instruments are derived from the standard neurological examination. For a detailed discussion of the psychometric properties of these instruments refer to Section 2.3.3. To maximise reproducibility, a single neurologist, Dr Jeremy Hobart, performed these assessments (Willoughby and Paty 1988). His level of intrarater reliability was high (n=32, intra class correlation coefficient = 0.92).

2. The Functional Independence Measure (FIM)
   Disability was assessed using the motor domain of the FIM. This 13 item scale assesses the level of functional independence in four subscales: selfcare, transfers, locomotion, and sphincter control. Each item is rated on a scale of
one to seven (one = complete dependence, seven = complete independence). The item scores are added to form a total score which can range from 13 to 91. For a detailed discussion of the psychometric properties of the FIM refer to Section 2.4.4. The scores were derived from patient interview by a single non-treating therapist (the researcher). This method of administration has been shown to be valid (Brousseau and Wolfson 1994). The intrarater reliability of the assessor was investigated and shown to be very high (intra class correlation coefficient for total motor domain = 0.99, refer to Appendix 9). In accordance with the guidelines, scores were consistently determined by actual performance of tasks on a daily basis rather than each individual’s optimum performance. This was important since the objective of rehabilitation is to impact on the patient’s performance in daily life.

3. The London Handicap Scale (LHS).

Handicap was assessed using the LHS. This self-report questionnaire assesses the level of disadvantage experienced by an individual in six dimensions: mobility, physical independence, occupation, social integration, orientation, and economic self-sufficiency. Each dimension is rated on a six point scale, arranged in order of increasing disadvantage. The raw scores are transformed to provide six dimension scores and an overall severity score which can range from 0 to 100 (where 0 = most disadvantaged, 100 = least disadvantaged). For a detailed discussion of the psychometric properties of the LHS refer to Section 2.5.4.
4. The Short Form 36 Health Survey Questionnaire (SF-36).

Health-related quality of life was assessed using the anglicised version of the SF-36 (Jenkinson et al 1993). This 36 item self-report questionnaire includes eight multi-item measures of functioning and well-being that represent physical and mental health status: physical functioning, role limitations physical or emotional health problems, social functioning, emotional well-being, pain, energy and vitality, and general health perceptions. All items are coded, summed and transformed onto a scale of 0 to 100, with 0 representing the poorest and 100 the optimal health on the scale. For a detailed discussion of the psychometric properties of the SF-36 refer to Section 2.6.7.

5. The 28-item General Health Questionnaire (GHQ).

Changes in emotional well-being were assessed by using the 28-item, self-report, GHQ. This version assesses four identifiable elements of emotional distress: depression, anxiety, social dysfunction and somatic complaints. Scores on each item are added to form four subscale scores and a total score which can range from 0 to 28, with higher scores indicating greater emotional distress. For a detailed discussion of the psychometric properties of the GHQ refer to Section 2.7.4.

The method of administration for the LHS, SF-36, and GHQ questionnaires was patient self-report. Whenever possible patients completed the questionnaires independently but where necessary (e.g. visual disturbance, difficulties with writing) physical assistance was provided by the researcher. No assistance was given in interpreting the questionnaires.
4.4.6 The rehabilitation service and programme

For purposes of comparison and replication it is important to detail aspects of the rehabilitation service utilised: the rehabilitation approach, the service process, the composition of staff, and the nature and duration of services offered. These are described below:

i) The rehabilitation approach

This was based on a model of comprehensive care (La Rocca et al 1994) which considers that rehabilitation management extends beyond symptomatic treatment, and emphasises the achievement of the best possible quality of life for the person within the limits of their disease (McGrath and Davis 1992; Schapiro and Langer 1994). While no two centres practice or deliver comprehensive care in an identical way (Thompson et al 1994), the literature demonstrates a commonality of practice between centres, with the key elements identified as:

1. a multidisciplinary team approach
2. interventions tailored to meet the individual’s needs, and
3. a patient centred functional goal setting approach.


Each of these key elements formed an integral part of the rehabilitation programme undertaken by the patients in the treatment group (for a more detailed review refer to Section 1.3.1).
ii) The service process

The routine clinical process undertaken at the NRU has been documented in some detail (Johnson and Thompson 1996; Rositter and Thompson 1995a; Freeman et al 1996a). It is schematically represented in Figure 4.2.

Fig 4.2 Schematic representation of the assessment, selection and rehabilitation service provided by the NRU, NHNN.

(reproduced with permission of the authors, Johnson and Thompson (1996))
iii) Staffing establishment

The staffing establishment is in line with the ideal compositions and staff:patient ratios recommended for neurological teams by the Association of British Neurologists (ABN Working Party Report 1992, p.13). The Consultant is a Neurologist who has a special interest in neurological disability. Junior medical staff include a Senior Registrar and a Senior House Officer. Sixteen whole time equivalent qualified Nurses and five Auxiliary Nurses are supervised by a full time Ward Sister with the Clinical Nurse Manager having overall responsibility for Nursing Services. Therapy staff include: four Physiotherapists (two senior II, one senior I, one superintendent III); four Occupational Therapists (two senior II, one senior I, one head III) and one Occupational Therapy Assistant; one point five Speech and Language Therapists (one head III, 0.5 senior I); one Clinical Psychologist; and one half time Social Worker.

iv) Nature and duration of services offered

The core generic services identified by the Royal College of Physicians in 1987 (Section 1.3.1, Table 1.2) are integral to the comprehensive multidisciplinary service offered. Additionally a comprehensive range of specialist services are available including: dietetics, urological services, orthotics and splinting services, communication aids assessment, and environmental controls service.

On average, all patients participate in two 45 minute sessions of physiotherapy and one session of occupational therapy per day. Speech and language therapy, social work and clinical psychology input is provided as required. The
nurses involvement, in ensuring carryover of newly acquired skills into
everyday function, is considered an essential component of the rehabilitation
service (Johnson 1995).

Integrated Care Pathways (Rossiter and Thompson 1995a,b), are used
routinely to document and monitor the clinical care provided for all MS patients
at the NRU. While this information is not utilised in this study, it is readily
available for purposes of replication or comparison.

4.4.7 Quality control of data input
To ensure accuracy of data input a detailed visual examination of all data was
routinely undertaken, and a random selection of 30% of the data was double
checked.

4.4.8 Statistical Analysis
Statistical advice was undertaken using SPSS Version 6.0 (SPSS Inc. 1993).
Descriptive statistics were used to describe the study population in terms of
demographic and disease characteristics, and initial baseline levels of
impairment (EDSS and FS), disability (FIM), handicap (LHS), health-related
quality of life (SF-36) and emotional well-being (GHQ) scores. Data was
routinely examined for distribution and homogeneity of variance.

Outcome scores were analysed in terms of change scores between the initial
and second assessment (Altman 1991; Ziebland 1994). Change scores were
computed by subtracting baseline scores from the six week assessment scores
(this calculation was reversed in computing EDSS and GHQ change scores where higher scores indicate deterioration). These scores were then summarised as a median/mean, standard deviation, and range. The difference between the treatment and control groups, in terms of change scores, was compared. The EDSS, the total FIM motor domain, and the individual subscale dimensions of the FIM and LHS were analysed using Wilcoxon Rank Sum Tests (Mann Whitney U). The total LHS change scores were analysed using Unpaired Student t-tests. For all estimates, a non-directional hypothesis (two-tailed test) was assumed, with a 1% level of significance to allow for multiple comparisons. In addition, changes were described in terms of the percentages of patients who had improved, deteriorated or remained the same in each of the measures.

Statistical significance does not assure clinical significance. It is therefore increasingly recommended that the 95% confidence interval (95% CI) and the effect size statistic should be used to supplement standard statistical testing to give a more complete and clinically relevant picture of health status change (Cohen 1977; Kaziz et al 1989; Deyo et al 1991; Daly et al 1991; Borenstein 1994; Mathews and Altman 1996). The confidence interval provides limits between which the population mean is likely to lie, revealing the precision of the estimate of the parameter measured (Gardner and Altman 1986). It provides information as to the amount of fluctuation that can be expected in the score due to measurement error (Ware 1994b), thereby estimating the uncertainty associated with a sample estimate (Bland 1994). The effect size transforms the score change into a standard unit of measurement which can be
used as a benchmark for interpreting the relative magnitude of change produced by the clinical intervention (Cohen 1977). It provides an indication as to whether a given change exceeds the measurement variability. Several different methods have been proposed to evaluate the effect size (Guyatt et al 1987; Kaziz et al 1989; Liang et al 1990). In this study the effect size was calculated according to Kaziz and colleagues (1989) where:

\[
\text{effect size} = \frac{\text{mean change}}{\text{standard deviation of the initial distribution of scores}}.
\]

An effect size of 1.0 is therefore equivalent to a change in one standard deviation of the sample. The criteria proposed by Cohen (1977) was used to interpret the effect size, where an effect size of 0.2 is small, 0.5 is moderate and 0.8 or greater is large.

### 4.5 Results

From a total sample of 279 consecutive patients assessed at the clinic over a 19 month period (May 1994 - January 1996), 112 were admitted for inpatient rehabilitation. Of these, 70 patients fulfilled the entry criteria and were recruited into the study. Forty one patients failed to meet the entry criteria (Table 4.2), and one person declined to participate in the study.

<table>
<thead>
<tr>
<th>Reasons for failing to meet criteria</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Required emergency admission as indicated on clinical grounds (e.g. pressure sores)</td>
<td>12</td>
</tr>
<tr>
<td>On steroids / in acute relapse</td>
<td>18</td>
</tr>
<tr>
<td>Within one month of relapse</td>
<td>2</td>
</tr>
<tr>
<td>Not definite MS</td>
<td>3</td>
</tr>
<tr>
<td>Severe cognitive disturbance</td>
<td>3</td>
</tr>
<tr>
<td>Other diseases which may interfere with outcome (peripheral demyelination x1; poorly controlled epilepsy x2)</td>
<td>3</td>
</tr>
</tbody>
</table>
Of the 70 patients entering the study (treatment = 34, control = 36), four withdrew prior to the second assessment. Details of these patients are summarised below:

- Case 4 (control) was a 46 year old female with secondary progressive MS (baseline scores: EDSS 7.5; FIM 50; LHS 57.4; GHQ 11). She withdrew from the study finding the questions too distressing.
- Case 10 (control) was a severely disabled 47 year old female with primary progressive MS (baseline scores: EDSS 9.0; FIM 13; LHS 48.9; GHQ 18). She withdrew from the study, finding the questions too distressing.
- Case 39 (treatment) was a 61 year old man with secondary progressive MS (baseline scores: EDSS 6.5; FIM 48; LHS 62.8; GHQ 4). He withdrew from the trial finding the burden of questions too great.
- Case 43 (treatment) was a 46 year old woman with secondary progressive MS (baseline scores EDSS 6.5; FIM 78; LHS 58.7; GHQ 2). She withdrew from the trial, finding the burdens of questions too great.

In summary, of the four patients who withdrew from the study, two were from the treatment group and two were controls. Two declined the second assessment finding the questions too distressing, while the other two found the burden of questions too great. Thus data from 66 patients was available for statistical analysis (treatment n = 32, control n = 34).
4.5.1 Sample: baseline characteristics

i) Demographic and diagnostic characteristics

Demographic and diagnostic characteristics of the two groups are summarised in Table 4.3. The two groups were well matched for all variables. It is notable that all patients were stratified to either the moderate or severe categories of disease severity (EDSS 5.0 - 9.5).

Table 4.3 Baseline demographic and diagnostic characteristics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Treatment group (n = 32)</th>
<th>Control group (n = 34)</th>
</tr>
</thead>
<tbody>
<tr>
<td>% female</td>
<td>66%</td>
<td>62%</td>
</tr>
<tr>
<td>Mean age (years), +/-sd, (range)</td>
<td>43.2, +/-10.77, (25 - 73)</td>
<td>44.6, +/-9.73, (25 - 61)</td>
</tr>
<tr>
<td>Disease pattern:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>secondary progressive</td>
<td>94%</td>
<td>88%</td>
</tr>
<tr>
<td>primary progressive</td>
<td>6%</td>
<td>12%</td>
</tr>
<tr>
<td>Mean years since first symptoms +/-sd, (range)</td>
<td>15.4, +/-7.6, (3 - 32)</td>
<td>15.6, +/-8.84, (3 - 35)</td>
</tr>
<tr>
<td>Mean years since diagnosis +/-sd, (range)</td>
<td>9.6, +/-6.4, (1.5 - 28)</td>
<td>11.4, +/-6.6, (2 - 23)</td>
</tr>
<tr>
<td>Relapse in past 6 months (%)</td>
<td>15%</td>
<td>6%</td>
</tr>
<tr>
<td>Mean number of months since relapse (if in last 6 months) +/-sd, (range)</td>
<td>4.2, +/-1.3, (3 - 6)</td>
<td>3.2, +/-1.4, (1.5 - 4)</td>
</tr>
<tr>
<td>Stratification categories (%) according to disease severity:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EDSS 0.0 - 4.5 (mild)</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>EDSS 5.0 - 6.5 (moderate)</td>
<td>53%</td>
<td>56%</td>
</tr>
<tr>
<td>EDSS 7.0 - 9.5 (severe)</td>
<td>47%</td>
<td>44%</td>
</tr>
</tbody>
</table>
**ii) Outcome measure scores**

Table 4.4 details the baseline values of all outcomes in the two groups.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Treatment group (n = 32)</th>
<th>Control group (n = 34)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FS: median, (range)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bladder/bowel</td>
<td>2, (0 - 6)</td>
<td>2, (0 - 6)</td>
</tr>
<tr>
<td>Brainstem</td>
<td>2, (0 - 4)</td>
<td>2, (0 - 4)</td>
</tr>
<tr>
<td>Cerebellar</td>
<td>2.5, (0 - 5)</td>
<td>2, (2 - 5)</td>
</tr>
<tr>
<td>Pyramidal</td>
<td>4, (2 - 5)</td>
<td>4, (2 - 5)</td>
</tr>
<tr>
<td>Sensory</td>
<td>2, (0 - 4)</td>
<td>3, (0 - 5)</td>
</tr>
<tr>
<td>Visual</td>
<td>1, (0 - 6)</td>
<td>2, (0 - 2)</td>
</tr>
<tr>
<td>Cerebral</td>
<td>0.5, (0 - 4)</td>
<td>1, (0 - 3)</td>
</tr>
<tr>
<td><strong>FIM motor domain (total)</strong></td>
<td>median, IQR 45.5 - 74.5, (13 - 87)</td>
<td>69.5, IQR 50 - 79, (18 - 84)</td>
</tr>
<tr>
<td><strong>LHS (total)</strong></td>
<td>mean, +/-sd, (range)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>61.5, +/-13, (32.6 - 87.2)</td>
<td>66.2, +/-8.74, (47.9 - 89.6)</td>
</tr>
<tr>
<td><strong>GHQ (total)</strong></td>
<td>median, IQR, (range)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5, IQR 4 - 9.8, (0 - 25)</td>
<td>7, IQR 2.7 - 13.5, (0 - 21)</td>
</tr>
<tr>
<td><strong>SF-36 dimensions</strong></td>
<td>mean, +/-sd, IQR, (range)</td>
<td></td>
</tr>
<tr>
<td>energy and vitality</td>
<td>38.9, +/-19.9, IQR 20 - 53.8, (5 - 80)</td>
<td>37.5, +/-20.2, IQR 25 - 50, (0 - 90)</td>
</tr>
<tr>
<td>health perceptions</td>
<td>40.8, +/-23.4, IQR 20.5 - 59.5, (5 - 100)</td>
<td>42.5, +/-27.9, IQR 18.8 - 63.2, (0 - 97)</td>
</tr>
<tr>
<td>mental health</td>
<td>62.1, +/-18.7, IQR 48 - 80, (32 - 100)</td>
<td>62.4, +/-19.5, IQR 47 - 81, (24 - 100)</td>
</tr>
<tr>
<td>pain</td>
<td>73.6, +/-26.9, IQR 55.6 - 100, (11.1 - 100)</td>
<td>62.1, +/-27.4, IQR 44.4 - 80.8, (11.1 - 100)</td>
</tr>
<tr>
<td>physical function</td>
<td>14.7, +/-15.1, IQR 1.2 - 23.8, (0 - 70)</td>
<td>11.8, +/-13.7, IQR 0 - 16.2, (0 - 45)</td>
</tr>
<tr>
<td>role limitation - mental</td>
<td>50, +/-44, IQR 0 - 100, (0 - 100)</td>
<td>43.1, +/-46.8, IQR 0 - 100, (0 - 100)</td>
</tr>
<tr>
<td>role limitation - physical</td>
<td>28.6, +/-38.6, IQR 0 - 50, range 0 - 100</td>
<td>8.8, +/-21.2, IQR 0 - 0, range 0 - 100</td>
</tr>
<tr>
<td>social function</td>
<td>51.3, +/-26.1, IQR 33.3 - 66.7, range 0-88.9</td>
<td>47.1, +/-21.7, IQR 33.3 - 66.7, range 0-88.9</td>
</tr>
</tbody>
</table>

IQR = interquartile range; SF-36 scores are transformed scores.
The two groups were well matched for all variables. The high scores on the EDSS in each group show that, at best, patients required intermittent or unilateral constant assistance to walk 100 metres, while at worst they were essentially bed bound. The severity of disease and high level of physical disability is confirmed by the relatively low scores on the FIM motor domain and the physical functioning dimensions of the SF-36.

Particularly severe problems with role limitations due to physical and mental problems and physical functioning were demonstrated by the low SF-36 scores. A marked floor effect was observed in these three dimensions where the lowest possible score of zero was scored by: 59% of the treatment group and 79% of controls for role limitations due to physical problems; 34% of the treatment group and 50% of controls for role limitations due to mental problems; and 25% of the treatment group and 35% of controls for physical functioning. Moderately severe problems were demonstrated in the dimensions of energy and vitality, social functioning, mental health and health perceptions. It is notable that compared to normative values for the British healthy population (Jenkinson et al 1993), both groups reported poorer levels of health in every single SF-36 dimension.

The GHQ scores confirm the presence of problems in the area of emotional health and well-being, although these appear less striking than the physical consequences. The GHQ Users Manual (Goldberg and Williams 1988) provides criteria for interpreting GHQ scores. A total GHQ score of greater than five points indicates the presence of emotional disturbance, 5-14 points being
moderately abnormal, 15-25 markedly abnormal, and 20-28 severely abnormal. This criterion has since been used by previous investigators studying neurological populations (e.g. Dalos et al 1982; Lykouras et al 1996). Using this criterion, 72% of the treatment group and 65% of controls scored equal to or greater than five points on the total GHQ score, indicating the presence of emotional disturbance. Of these, 19% of the treatment group and 23% of controls scored equal to or greater than 15 points indicating a "markedly abnormal" level of emotional disturbance. Vickrey and colleagues (1995) used a "cut-off" score of 66 on the mental health scale of the SF-36 to divide their population into those who had scores associated with depressive symptoms and those who did not. Using this criterion, 59% of the treatment group and 64% of controls demonstrated depressive symptoms.

4.5.2 The rehabilitation intervention

i) Rehabilitation programme

All patients in the treatment group received an individualised, goal-oriented programme (as detailed in Section 4.4.6) addressing a wide range of areas (Table 4.5) for an average of 20 days (sd 3, range 17-31). All patients had medical, nursing, occupational therapy and physiotherapy input. Eighty five per cent of patients were assessed by the neuropsychologist, 64% by the speech and language therapist, and 48% by the social worker. Consultation was obtained as required from psychiatric, urological and dietetic services.
Table 4.5 Broad areas addressed in the rehabilitation programme

<table>
<thead>
<tr>
<th>Broad areas addressed</th>
<th>% patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Provision of home exercise programme</td>
<td>97</td>
</tr>
<tr>
<td>Re-education posture and balance</td>
<td>94</td>
</tr>
<tr>
<td>Fatigue management</td>
<td>90</td>
</tr>
<tr>
<td>Personal activities of daily living</td>
<td>84</td>
</tr>
<tr>
<td>Cognitive assessment</td>
<td>83</td>
</tr>
<tr>
<td>Bladder management</td>
<td>80</td>
</tr>
<tr>
<td>Posture and seating</td>
<td>74</td>
</tr>
<tr>
<td>Advice to carers</td>
<td>71</td>
</tr>
<tr>
<td>Gait re-education</td>
<td>68</td>
</tr>
<tr>
<td>Leisure pursuits</td>
<td>64</td>
</tr>
<tr>
<td>Wheelchair skills</td>
<td>61</td>
</tr>
<tr>
<td>Domestic activities of daily living</td>
<td>61</td>
</tr>
<tr>
<td>Tone management</td>
<td>55</td>
</tr>
<tr>
<td>Bowel management</td>
<td>52</td>
</tr>
<tr>
<td>Communication skills</td>
<td>48</td>
</tr>
<tr>
<td>Emotional assessment and counselling</td>
<td>42</td>
</tr>
<tr>
<td>Swallowing assessment and management</td>
<td>35</td>
</tr>
<tr>
<td>Drug management</td>
<td>32</td>
</tr>
<tr>
<td>Work skills</td>
<td>29</td>
</tr>
<tr>
<td>Sexual function</td>
<td>3</td>
</tr>
</tbody>
</table>

**ii) Drug management**

On initial assessment 91% of patients in the treatment group, and 82% of the control group were taking some form of prescribed medication. By the second assessment six weeks later, this had increased to 100% and 86% respectively (Table 4.6). Only minor changes to the overall drug management of patients in either group were noted. The most common changes were in the areas of bladder management and spasticity.
Table 4.6

Drug management of study patients (n = 34 control, n = 32 treatment)

<table>
<thead>
<tr>
<th>Drug category</th>
<th>Control group</th>
<th>Control group</th>
<th>Rx group</th>
<th>Rx group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0 weeks</td>
<td>6 weeks</td>
<td>0 weeks</td>
<td>6 weeks</td>
</tr>
<tr>
<td>Immunosuppressants</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Anti-spasticity agents</td>
<td>11</td>
<td>12</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>Anti-depressants</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Anti-tremor agents</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Fatigue management</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Bowel management</td>
<td>9</td>
<td>10</td>
<td>20</td>
<td>21</td>
</tr>
<tr>
<td>Bladder management</td>
<td>9</td>
<td>10</td>
<td>6</td>
<td>13</td>
</tr>
<tr>
<td>Analgesics</td>
<td>7</td>
<td>8</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Other</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>15</td>
</tr>
</tbody>
</table>

Rx group = Treatment group

4.5.3 Outcome measurement change scores

The primary outcome criteria were comparisons between group changes in the zero and six week assessment scores.

i) Impairment

The neurological status, as measured by the EDSS and FS change scores, did not change for either the treatment or control group over the six week period (Table 4.7). There were no significant differences between the two groups in either the EDSS change score (Wilcoxon Rank Sum Test; p = 0.4202) or any of the FS Scale change scores. This lack of change was confirmed by the very small effect sizes in all change scores for both groups. The 95% confidence intervals for the impairment change scores demonstrate that the magnitude of change was negligible, with change scores clustering closely around zero.
<table>
<thead>
<tr>
<th></th>
<th>Treatment group</th>
<th>Effect size</th>
<th>Mean (95% CI)</th>
<th>Control group</th>
<th>Effect size</th>
<th>Mean (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Change score</td>
<td>Treatment</td>
<td></td>
<td></td>
<td>Treatment</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(range)</td>
<td>group</td>
<td></td>
<td>(range)</td>
<td>group</td>
<td></td>
</tr>
<tr>
<td>EDSS</td>
<td>0 (-1 - +1)</td>
<td>0.06</td>
<td>0.06 (-0.08, 0.21)</td>
<td>0 (-1 - +0.5)</td>
<td>0.1</td>
<td>0.10 (-0.01, 0.21)</td>
</tr>
<tr>
<td>Functional Systems:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pyramidal</td>
<td>0 (-1 - +1)</td>
<td>0.12</td>
<td>0.09 (-0.12, 0.32)</td>
<td>0 (-1 - +1)</td>
<td>0.13</td>
<td>0.15 (-0.03, 0.33)</td>
</tr>
<tr>
<td>Cerebellar</td>
<td>0 (-3 - +1)</td>
<td>-0.2</td>
<td>-0.32 (-0.74, 0.09)</td>
<td>0 (-2 - +2)</td>
<td>0.09</td>
<td>0.03 (-0.27, 0.33)</td>
</tr>
<tr>
<td>Brainstem</td>
<td>0 (-2 - +1)</td>
<td>0</td>
<td>0 (-0.28, 0.28)</td>
<td>0 (-2 - +2)</td>
<td>0.09</td>
<td>0.06 (-0.24, 0.34)</td>
</tr>
<tr>
<td>Sensory</td>
<td>0 (-2 - +3)</td>
<td>0.15</td>
<td>0.23 (-0.23, 0.68)</td>
<td>0 (-2 - +2)</td>
<td>0.08</td>
<td>0.61 (-0.33, 0.45)</td>
</tr>
<tr>
<td>Bowel/bladder</td>
<td>0 (-2 - +3)</td>
<td>0</td>
<td>-0.32 (-0.38, 0.32)</td>
<td>0 (-1 - +1)</td>
<td>0.06</td>
<td>0.09 (-0.15, 0.33)</td>
</tr>
<tr>
<td>Visual</td>
<td>0 (-2 - +1)</td>
<td>-0.13</td>
<td>-0.16 (-0.45, 0.12)</td>
<td>0 (-2 - +2)</td>
<td>-0.07</td>
<td>-0.12 (-0.42, 0.18)</td>
</tr>
<tr>
<td>Cerebral</td>
<td>0 (-2 - +3)</td>
<td>0.16</td>
<td>0.16 (-0.25, 0.57)</td>
<td>0 (-2 - +1)</td>
<td>-0.09</td>
<td>-0.12 (-0.40, 0.10)</td>
</tr>
</tbody>
</table>

Positive changes imply improvement


**ii) Disability**

Table 4.8 details the disability change scores as measured by the total FIM motor domain and its four subscales.

**Table 4.8 FIM change scores for treatment and control groups**

<table>
<thead>
<tr>
<th></th>
<th>Treatment group (n = 32)</th>
<th>Control group (n = 34)</th>
<th>p value *</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>median</td>
<td>IQR</td>
<td>range</td>
</tr>
<tr>
<td>FIM total score</td>
<td>+ 4.0</td>
<td>-0.8-+8.5</td>
<td>-10-+19</td>
</tr>
<tr>
<td>Sub-scales:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self care</td>
<td>+ 1.5</td>
<td>0+-5</td>
<td>-5+-9</td>
</tr>
<tr>
<td>Transfers</td>
<td>0</td>
<td>0-+2</td>
<td>-4-+6</td>
</tr>
<tr>
<td>Sphincter</td>
<td>0</td>
<td>0-+2</td>
<td>-1-+6</td>
</tr>
<tr>
<td>Locomotion: Walking</td>
<td>+ 1.0</td>
<td>-1-+2</td>
<td>-0.6-+6</td>
</tr>
<tr>
<td>treatment n = 17;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>control n = 20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>wheelchair</td>
<td>0</td>
<td>0-+1</td>
<td>-3-+5</td>
</tr>
<tr>
<td>treatment n = 15;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>control n = 14</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Wilcoxon Rank Sum W Test; Positive changes imply improvement

A statistically significant difference between the two groups (p < 0.001) in terms of overall motor domain change scores was demonstrated (Figure 4.3).
Figure 4.3 Boxplot chart showing the difference in disability change scores between the treatment and control groups

Boxplot represents the median, interquartile range and range of scores,

Analysis of the subscale change scores demonstrated statistically significant differences between the treatment and control groups in three of the four subscales: self care ($p < 0.0001$); transfers ($p < 0.001$); and sphincter control ($p < 0.001$). In terms of the locomotion subscale, patients were sub-grouped according to whether they walked (treatment $n = 17$, control $n = 20$) or used a wheelchair (treatment $n = 15$, control $n = 14$). No statistically significant difference was shown between the treatment and control groups for those who were walking ($p = 0.3814$), or for the wheelchair users ($p = 0.0315$).
Effect sizes and 95% confidence intervals were calculated for the total disability scores. The relatively large variance of the initial distribution of scores meant that the gains in raw scores translated into small effect sizes. In terms of the magnitude of functional change, the treatment group made a small improvement (effect size +0.21, mean +3.9, 95% CI +1.76, +6.12), in contrast to the control group who made a negligible decline (effect size -0.16, mean -2.97, 95% CI -4.76, -1.18).

The FIM motor domain scores were also used to describe the changes which occurred in terms of the percentage of patients who improved, deteriorated or remained the same in each group. In this regard, 72% of patients in the treatment group improved their overall level of disability, 3% stayed the same and 25% deteriorated. In contrast, 29% of patients in the control group improved their overall level of disability, 9% stayed the same and 62% deteriorated.

**iii) Handicap**

Table 4.9 details the handicap change scores as measured by the LHS. A statistically significant difference between the two groups in terms of the changes made in the overall level of handicap was demonstrated ($p < 0.01$). Analysis of the individual handicap dimension change scores demonstrated that these differences did not reach statistical significance in any of the six handicap dimensions. No individual dimension of the LHS appeared to contribute more to the total handicap severity score in either group.
Table 4.9 LHS change scores for treatment and control groups

<table>
<thead>
<tr>
<th></th>
<th>Treatment group (n = 32)</th>
<th>Control group (n = 4)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean +/-sd</td>
<td>range</td>
<td>mean +/-sd</td>
</tr>
<tr>
<td>LHS total score</td>
<td>+2.9 +/- 8.9</td>
<td>-9.2 - +36.2</td>
<td>-2.7 +/- 8.6</td>
</tr>
<tr>
<td>Individual dimension</td>
<td>median</td>
<td>IQR</td>
<td>median</td>
</tr>
<tr>
<td>Mobility</td>
<td>0</td>
<td>0 - +2.9</td>
<td>0</td>
</tr>
<tr>
<td>Physical independence</td>
<td>0</td>
<td>0 - +1.5</td>
<td>0</td>
</tr>
<tr>
<td>Occupation</td>
<td>0</td>
<td>0 - +0.7</td>
<td>0</td>
</tr>
<tr>
<td>Social interaction</td>
<td>0</td>
<td>-2.5 +2</td>
<td>0</td>
</tr>
<tr>
<td>Orientation</td>
<td>0</td>
<td>0 - 0</td>
<td>0</td>
</tr>
<tr>
<td>Economic self sufficiency</td>
<td>0</td>
<td>0 - 0</td>
<td>0</td>
</tr>
</tbody>
</table>

Positive scores imply improvement; * Unpaired t-tests; ** Wilcoxon Rank Sum Tests
On average, the absolute changes in the total handicap scores were relatively small, being +2.9 points in the treatment group and -2.7 points in the control group (Figure 4.4).

**Figure 4.4** Boxplot chart showing the difference in handicap change scores between the treatment and control groups.

Effect sizes and 95% confidence intervals were calculated for the total handicap change scores. The relatively large variance of the initial distribution of scores meant that the gains in raw scores translated into small effect sizes. In terms of the magnitude of change in overall handicap, the treatment group made a small improvement (effect size +0.23, mean +2.76, 95% CI -0.44, +5.96) in contrast to the control group who made a small decline (effect size -0.27, mean -2.71, 95% CI -5.73, +0.29).
In terms of the percentage of patients who improved, deteriorated or remained the same in each group, 53% of the treatment group improved their total handicap score, 3% remained the same and 44% deteriorated. In contrast, 23% of the control group improved their total score, 12% stayed the same and 65% deteriorated.

**iv) Health-related Quality Of Life**

Table 4.10 details the change scores for the eight dimensions of the SF-36. Descriptive statistics revealed no notable differences between the two groups in any of the dimensions. It is notable that the range and standard deviations of the change scores are all large relative to the mean, suggesting considerable variability in the changes reported by individual subjects.

The lack of change was confirmed by the very small effect sizes for all SF-36 dimensions in both groups (Table 4.11). The wide 95% confidence intervals for each of the change scores confirm the large variability in change scores.

The SF-36 Manual and Interpretation Guide (Ware 1993b) provides criteria for interpreting change, defining a two point change as "very small"; a five point difference as "clinically and socially relevant change"; a 10 point change as "moderate"; and a 20 point change as "large". Based on this criterion, the only dimension to demonstrate clinically and socially relevant change within the six week assessment period was the role limitations due to mental problems. This was demonstrated by the treatment group.
Table 4.10 Profile of SF-36 change scores for treatment and control groups

<table>
<thead>
<tr>
<th>SF-36 dimensions</th>
<th>Treatment Group (n = 32)</th>
<th>Control Group (n = 34)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean +/-sd</td>
<td>range</td>
</tr>
<tr>
<td>energy and vitality</td>
<td>+3.9 +/-21.3</td>
<td>-50 - +70</td>
</tr>
<tr>
<td>health perceptions</td>
<td>+1.9 +/-19.3</td>
<td>-37 - +58</td>
</tr>
<tr>
<td>mental health</td>
<td>+4.2 +/-16.9</td>
<td>-28 - +60</td>
</tr>
<tr>
<td>pain</td>
<td>-4.2 +/-26.3</td>
<td>-77.8 - +55.6</td>
</tr>
<tr>
<td>physical function</td>
<td>+2.0 +/-7.3</td>
<td>-10 - +20</td>
</tr>
<tr>
<td>role limitation - mental</td>
<td>+5.2 +/-45.7</td>
<td>-100 - 100</td>
</tr>
<tr>
<td>role limitation - physical</td>
<td>-3.8 +/-49.2</td>
<td>-100 - +100</td>
</tr>
<tr>
<td>social function</td>
<td>+1.7 +/-25.8</td>
<td>-44 - +66.7</td>
</tr>
</tbody>
</table>

Positive scores imply improvement
<table>
<thead>
<tr>
<th>SF-36 dimensions</th>
<th>Treatment group (n = 32)</th>
<th>Control group (n = 34)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>effect size</td>
<td>mean change</td>
</tr>
<tr>
<td>energy and vitality</td>
<td>0.19</td>
<td>+3.9</td>
</tr>
<tr>
<td>health perceptions</td>
<td>0.08</td>
<td>+1.9</td>
</tr>
<tr>
<td>mental health</td>
<td>0.22</td>
<td>+4.2</td>
</tr>
<tr>
<td>pain</td>
<td>-0.16</td>
<td>-4.2</td>
</tr>
<tr>
<td>physical function</td>
<td>0.13</td>
<td>+2.0</td>
</tr>
<tr>
<td>role limitation - mental</td>
<td>0.12</td>
<td>+5.2</td>
</tr>
<tr>
<td>role limitation - physical</td>
<td>0.09</td>
<td>-3.8</td>
</tr>
<tr>
<td>social function</td>
<td>0.06</td>
<td>+1.7</td>
</tr>
</tbody>
</table>

Cl = confidence interval; Positive scores imply improvement
With regard to the percentages of patients in each group who improved, deteriorated or remained the same, the frequency distribution showed that, compared to the controls, more patients in the treatment group reported improved levels of health in all but one dimension (social functioning). In this dimension the percentages were similar in both groups.

v) Emotional well-being

Table 4.12 details the change scores for the total GHQ and each of its four subscales.

<table>
<thead>
<tr>
<th></th>
<th>Treatment group (n = 32)</th>
<th>Control group (n = 34)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>median</td>
<td>IQR</td>
</tr>
<tr>
<td>GHQ total score</td>
<td>+1</td>
<td>-1.8 - +5.8</td>
</tr>
<tr>
<td>Subscale scores</td>
<td></td>
<td></td>
</tr>
<tr>
<td>anxiety and insomnia</td>
<td>0</td>
<td>0 - +1</td>
</tr>
<tr>
<td>somatic</td>
<td>0</td>
<td>-1 - +2</td>
</tr>
<tr>
<td>social dysfunction</td>
<td>0</td>
<td>-0.8 - +2</td>
</tr>
<tr>
<td>depression</td>
<td>0</td>
<td>0 - +0.8</td>
</tr>
</tbody>
</table>

Positive scores imply improvement
There was no statistically significant difference between the two groups in terms of total GHQ scores (Wilcoxon Rank Sum Test, p = 0.6528). Similarly, descriptive statistics revealed no discernible differences between the two groups in any of the subscale change scores. Overall, the GHQ change scores were small in both groups, with a narrow interquartile range (for both total and subscale scores) and narrow 95% confidence intervals. The majority of change scores tended to cluster around zero points indicating little change in the levels of emotional well-being in either group. Although the absolute improvement in the GHQ total scores was only one point, this translated into a small effect size in the treatment group because of the small standard deviation of the initial distribution. The effect sizes revealed a small improvement in overall level of emotional status in the treatment group (effect size +0.29; mean +1.72, 95% CI -0.66, +4.09) and a negligible improvement in the controls (effect size +0.15, mean +1, 95% CI -1.48, +3.53).

The percentage of patients who scored greater than five points (indicating the presence of emotional disturbance) decreased in both groups over the study period (at recruitment - treatment group 72%, control group 65%; at six weeks - 56% and 53% respectively). The frequency distribution of change scores showed that approximately the same percentage of patients in each group reported their emotional well-being to have either improved, deteriorated or remained the same at each assessment. This was the case for the total GHQ scores as well as each of its subscales.
4.5.4 Self-reported relapses

Two patients in the treatment group reported a relapse during the six week study period. This was confirmed by the assessing neurologist in one case. In addition, one person in the treatment group underwent an emergency appendectomy during the fourth week of the study, two days post discharge from the rehabilitation unit.

4.6 Discussion

4.6.1 Findings

The results of this study demonstrate that in the context of unchanging neurological impairment, both disability and handicap were reduced in the treatment group compared to the wait-list control group. No notable alteration in health-related quality of life or emotional well-being was evident in either group. The null hypothesis is thereby rejected in part. Patients with progressive MS participating in intensive multidisciplinary inpatient rehabilitation demonstrated better clinical outcome than those in the control group, in terms of disability and handicap.

While it is generally agreed that inpatient rehabilitation has the potential to increase the functional independence of patients, few studies have provided objective evidence to substantiate this assumption. This is in part, due to the complexity of designing studies in this area (Colville 1996; Hobart and Thompson 1997), over and above those already recognised in drug trials (Noseworthy 1996). Developing a methodologically sound study design, which remains clinically relevant and practical, is essential for results to be
convincing, but has proven difficult to achieve. This study’s approach has been to design a stratified randomised wait-list controlled design, with a relatively homogenous population of patients in the progressive phase of the disease, to enable more confident interpretation of the results.

The inclusion of a control group enables one to be reasonably confident that these changes were a direct result of the rehabilitation intervention rather than changes which may have occurred by chance. Randomisation and stratification ensured that the two groups were well matched for all baseline variables. The results would have been strengthened had the study design included a sham intervention as a control. Given the modest benefits resulting from the comprehensive rehabilitation programme, the changes in environment and attention received by the treatment group may have been an important factor in the observed improvements.

It is often suggested that functional improvements gained during a course of rehabilitation occur predominately as a direct result of spontaneous neurological recovery. To date the evidence to disprove this argument has been scant. There is little doubt that functional improvement in MS is multifactorial. Contributing factors may include spontaneous neurological recovery, and pharmacological intervention. It was attempted, as far as possible, to reduce the likelihood of improvement due to these factors by the inclusion and exclusion criteria. This was clearly not entirely possible in a disease whose course is unstable and unpredictable in nature, however, it was limited to a reasonable degree.
A six week time period for this study was chosen for two main reasons: (1) it is soon enough after discharge to be reasonably confident that the change is due to the intervention and not to other factors and; (2) ethical dilemmas arise in withholding therapies, which are thought to be effective, on the basis of random assignment to a treatment group in a controlled trial. The six week waiting time was similar to the usual waiting list period and therefore was acceptable to both patients and clinicians. As a consequence, bias in terms of patient recruitment was negligible, with only one person refusing to enter the study.

Alterations in drug management did not have a notable effect on outcome. Only minor changes in medication occurred in either of the two groups. These changes were predominantly in the area of bladder management, where seven patients in the treatment group were commenced on oxybutanin compared to one patient in the control group. Despite this, changes on the FIM sphincter subscale, where this drug would be expected to most influence outcome, did not show the greatest improvement, and the control group declined on this subscale in a similar manner to all the other subscales.

The EDSS and individual FS scales showed no change as a result of the rehabilitation input. This is not surprising for a number of reasons. Firstly, disease activity is not expected to change as a direct result of rehabilitation in patients in the progressive phase of MS. This is in contrast to clinical trials of therapeutic pharmacological agents where the primary aim of intervention is to influence aspects such as relapse rate (e.g. IFNB Study Group 1995; Johnson et al 1995; Jacobs et al 1996). Secondly, all patients in this study scored in the
upper section of the EDSS scale, in which there is acknowledged to be a major emphasis on locomotion (Noseworthy 1994a). It is noteworthy that the only aspect of the FIM which did not show significant change was mobility. The fact that changes in the EDSS were not significant does not, however, mean that individual patients scores did not change on this measure. Other studies, particularly those who have involved patients with relapsing-remitting MS (where there is the likelihood of spontaneous neurological recovery and consequent improvement in impairment), have shown changes on the EDSS (Kidd et al 1995; Aisen et al 1996). It is also possible to register improvement on the EDSS without changes in impairment. For example, if patients improve their transfers, or reduce the number of aids required for walking, they will show a change on this mobility dominated scale (Kidd et al 1995).

The results demonstrate that rehabilitation is effective in reducing the overall level of disability. These findings are in agreement with all of the previous outcome studies investigating comprehensive inpatient rehabilitation in MS. Of these studies, the majority restricted their analysis to total disability scores (Reding et al 1987; Carey et al 1988; Francabandera et al 1988; Kidd et al 1995; Cendrowski et al 1996; Vaney et al 1996a; Kidd and Thompson 1997, in press). A total score, however, does not show where changes have occurred and where disabilities remain. The use of subscale or individual item analysis can improve score interpretability. Analysis of the FIM subscale scores demonstrated that three of the four (transfers, self care activities, and sphincter control) significantly improved as a result of the rehabilitation programme. These results are in agreement with two of the three earlier studies which
analysed subscale scores (Feigenson et al 1981; Greenspun et al 1987). They contrast with the results of Aisen and colleagues (1996) who demonstrated significant improvements in sphincter control, self care and locomotor status, but not in transfer ability.

In the FIM locomotion subscale no statistically significant differences were evident between the two groups. This is not surprising. In order to demonstrate improvement on this subscale, relatively large reductions are required in either the amount of assistance or the distances negotiated. For example, to shift from level five (where a quarter of both the treatment and control patients were initially scored) to six on the FIM locomotion item, the patient must improve from being able to walk, or manoeuvre a chair, a distance of 17 metres to a minimum distance of 50 metres. Clinical experience suggests that such marked differences are unlikely to occur in this group of moderate to severely disabled patients where fatigue, muscle weakness, and spasticity are often major problems. While this finding supports the results of some studies (Feigenson et al 1981; Fuller et al 1996), it contradicts others (Greenspun et al 1987, Aisen et al 1996). Meaningful comparisons, however, are difficult to make due to the different sample populations and methodologies utilised in each of these studies. Furthermore, the FIM locomotion item does not take into account the quality of walking, a feature which is often of key interest to the therapist. It is suggested that research investigating this aspect of disability should use instruments designed specifically to measure mobility (e.g. the Rivermead Motor Index and the 10 metre timed walk) which have been shown to detect the magnitude of changes which can be realistically achieved in MS patients.
(Vaney et al 1996b), and which offer additional information about the pattern and speed of gait.

In terms of change in the overall level of handicap a statistically significant difference was demonstrated between the two groups. To date, only two previous studies have used standardised measures to investigate the impact of comprehensive MS inpatient rehabilitation on overall levels of handicap (Kidd et al 1995; Kidd and Thompson 1997, in press). Both of these uncontrolled, before-after-studies, determined the percentage of patients who had improved their handicap scores between admission and discharge (44% and 47% respectively). The results of this study was similar, demonstrating that 53% of patients in the treatment group had improved their overall handicap score at the second assessment.

No statistically significant differences between the two groups were shown in the individual handicap dimensions. The interquartile ranges do suggest however that in the dimensions of mobility, physical independence and occupation, larger gains were made by the treatment group compared to the control group. These are the areas where changes are most likely to occur following a short period of inpatient rehabilitation (Roy et al 1992). In contrast, changes in orientation, social integration or economic self sufficiency are less likely to be observed, especially within a six week assessment period. For example, with respect to economic self-sufficiency, 22% of the treatment group and 21% of controls scored at level four on the LHS indicating reliance on state benefits for extra expenses resulting from ill-health. It is unlikely that
rehabilitation will reduce dependence on state benefits in this population where 87% of patients in the treatment group and 80% of controls were unemployed.

Differences between the two groups were more significant in disability than handicap change scores. This may relate to a number of issues. Firstly that inpatient rehabilitation programmes tend to focus more heavily on disability rather than handicap issues (Davis et al 1992). Secondly, that changes are best measured closest to the level of intervention (Harwood and Ebrahim 1995), which in this case was disability. Thirdly, that the six week time frame of this study was too short for the provision and installation of many of the adaptations recommended by the rehabilitation team (e.g. the provision of aids and adaptations to improve access [rails, ramps, stairlift] and increase autonomy [wheelchair, environmental controls]). Fourthly, that the LHS is less responsive than the FIM. Longer term follow up once the patient has returned to their home environment is required to determine whether handicap is further reduced following discharge from the rehabilitation unit. The longitudinal study undertaken in parallel with this randomised controlled trial, was designed to address some of these issues (Experiment 2, Chapter 5).

Statistical significance does not assure clinical significance. Currently no criteria are available for interpreting what represents clinically meaningful change on either the FIM or the LHS. This makes meaningful interpretation of the raw change scores difficult. This problem was addressed, in part, by calculating the effect size statistic for the change scores. According to the criterion proposed by Cohen (where an effect size of 0.2 is small, 0.5 is
moderate and one of 0.8 or greater is large) the results demonstrated that the changes in disability and handicap were not only statistically significant but were also clinically meaningful in terms of general effect size benchmarks (Cohen 1977). The treatment group demonstrated small improvements in both disability and handicap, in contrast to the control group who declined in both domains. This decline was negligible in disability and small in handicap. Although small, the magnitude of these effect sizes are in line with clinical expectations for this population. The degree of clinical change observed in disability and handicap tend to be modest and undramatic in chronic progressive diseases such as MS (Harwood and Ebrahim 1995; Harwood et al 1996).

In this study the improvements in disability and handicap were independent of improvements in the patients neurological status. From a clinical perspective these results are unsurprising (Greenwood 1992; Edwards 1996; Wade 1996). Rehabilitation interventions, in chronic and progressive conditions, often specifically aim to educate the patient to use and refine compensation strategies to enable activities to be performed as efficiently as possible (Pope 1992; Mertin and Paeth 1994; Edwards 1996). These complex relationships between different dimensions of health have also been documented in other areas of MS research (e.g. MRI studies; Losseff et al 1996). They remain poorly understood (Wade 1996).

It is notable that there was some deterioration in disability and handicap in the control group over the six week assessment period. Other studies have also
reported deterioration in the control group over a similar time interval in patients with MS (Fuller et al 1996) and Stroke (Wade et al 1992). This may, in part, reflect a referral bias. In general, patients are referred for rehabilitation assessment because their functional status is deteriorating. The presenting problems are often complex and interrelated in their cause and effect (Johnson and Thompson 1996). Difficulty with one task may impact upon others, compounding the initial problem, and causing a downward spiral in a range of functional activities. Without any intervention it would appear that this process of deterioration continued in the control group.

Both groups in this study reported relatively poor levels of HRQOL in all dimensions of the SF-36. This highlights the wide ranging aspects of everyday life which are commonly affected as a consequence of MS. It supports the findings of previous studies which demonstrated lower HRQOL in patients with MS relative to: the general population (Vickrey et al 1995; Freeman et al 1996b; Aronson 1997); patients with inflammatory bowel disease and rheumatoid arthritis (Rudick et al 1992); stroke patients participating in an inpatient rehabilitation programme (Borgel et al 1992); and patients with epilepsy or diabetes (Hermann et al 1996).

Limited change was demonstrated in HRQOL by either group. There are several possibilities as to why this occurred: (1) changes in HRQOL did not occur; (2) the sample size was too small to detect changes; (3) the SF-36 was unresponsive in detecting any changes which may have taken place. Interpretation of changes scores must be made with consideration to the
instrument’s measurement properties (Fitzpatrick 1994). In terms of responsiveness this proved difficult since responsiveness data is not available on the SF-36 for the MS population. It is impossible to know which of these explanations may be correct. However, since 1994 (following commencement of this study) a growing body of evidence has become available demonstrating that marked floor effects are seen on the SF-36 in a variety of patient groups. These groups include: elderly patients (Hill and Harries 1994); patients with chronic physical conditions (Mawson 1995); and MS patients with moderate to severe physical disability (Freeman et al 1996b). Such marked floor effects were demonstrated by both the treatment and control groups in this study. This is likely to mean that the lower range of the scale, where the majority of the population is cited, was too limited to enable small but possibly clinically significant changes to be registered by the SF-36 (Bindman et al 1990).

Both groups demonstrated moderately high GHQ scores and relatively low scores on the SF-36 emotional health dimensions. This supports previous findings that emotional disturbance is a common experience in MS (e.g. Whitlock and Siskind 1980; Dalos et al 1983; Minden and Schiffer 1990). The percentage of patients identified as reporting some degree of emotional disturbance (72% of the treatment group and 65% of controls) is considerably higher than in other studies where prevalence levels were estimated at between 40% and 54% (Joffe et al 1987; Rabins et al 1986; Vickrey et al 1995). However, these findings are similar to those demonstrated in previous studies at the NRU in which 64% (n = 38, median EDSS 7.5, Langdon and Thompson 1995) and 70% (n = 50, median EDSS = 6.5, Freeman et al 1996b) of the MS
population demonstrated emotional disturbance. A number of explanations are possible. For example, that the presence of somatic complaints may have overinflated the GHQ scores (Wade 1992). Alternatively, it has been demonstrated that levels of emotional disturbance may be higher in MS when a more advanced level of physical disability has been reached (McLellan et al 1989; Devins et al 1993), as has found to be the case in other chronic conditions such as rheumatoid arthritis (Devins et al 1993), Parkinson's disease (Peto et al 1995) and ischaemic stroke (Kwa et al 1996).

No significant differences in GHQ change scores were noted between the treatment or control group. This suggests that a short period of multidisciplinary inpatient rehabilitation was not associated with improved levels of emotional distress. This contrasts with the findings of Jonsson and colleagues (1996) who found a significant decrease in aspects of emotional well-being following inpatient rehabilitation, as measured by the Beck Depression Inventory. It also contrasts with the perceptions of clinicians, patients and carers who often report subjective improvements in mood during rehabilitation. It is difficult to know why the results of this study differ from those of Jonsson and with general clinical perceptions. One possibility is that the GHQ does not measure some of the aspects of emotional well-being which change as a result of the rehabilitation process. Another possibility is that the improved levels of emotional well-being reported by 65% of patients in the control group (compared to 72% of the treatment group) may have influenced the statistical findings. The improvements reported by the control group may be partly associated with the timing of the second assessment, which was undertaken on
the day of admission to the rehabilitation unit. Perhaps the very prospect of immediate rehabilitation may have resulted in a feeling of optimism and an improved level of emotional well-being in these patients. Further studies are required to investigate this area in more detail.

4.6.2 Evaluation of methodology

i) Study design

Double-blind randomised placebo controlled studies are considered the gold-standard methodology for evaluating therapeutic effectiveness in MS (Hobart and Thompson 1997). This is difficult to achieve in practical terms in rehabilitation research. The design used was practical and overcame many methodological objections. Some however were not resolved, notably: the inability to blind the assessors to group allocation; to control for non-specific factors (e.g. changes in environment or attention); or for the placebo effect. It is recognised that these factors may lead to an over-estimation of treatment effects through a type 1 error (Ottenbacher and Jannell 1993; Noseworthy 1994b).

The main possible source of bias was the non-blinded assessors. This was minimised in a number of ways. Firstly non-treating assessors were used. Secondly, the time interval between the two assessments was relatively long (six weeks). Thirdly, the assessors had no further access to initial scores. Fourthly, a self-report method of data collection was used to measure four of the five domains, namely: disability, handicap, HRQOL and emotional well-being. This was perhaps of most relevance to the latter three which were all
patient administered, but of some relevance to disability which was scored from patient interview and therefore open to some interpretation by the rater.

It has been recommended that when blinding is impossible it is desirable to use more than one source of data collection, since when the results of different measurement methods converge, conclusions are strengthened, and when they diverge, otherwise premature interpretation can be avoided (Fuhrer 1987). In addition to the researchers disability scores, the patients were independently rated, as part of standard clinical procedure, by their treating multidisciplinary team who were blinded to group allocation. These scores were analysed (Appendix 10). Correlation of the treating team scores with those of the research assessor showed a very high correlation (Spearman's correlation coefficient = 0.95 - 0.96), which supports the validity of this method of disability scoring.

The patients were not blind to the intervention since there was no direct placebo treatment. It is recognised that, particularly when not blinded, patients may over-report or under-report their symptoms to “please the interviewer”. Had this been the case, one might have expected scores on all measures to consistently either improve or decline according to group allocation. This did not happen. In addition, the very high correlation of the treating team’s disability scores with the patient self report disability scores goes some way to demonstrate that the levels reported by the patients represented their current functional status.
This study did not include a placebo-control group. Substantial improvements on this aspect of the study design, however, may prove difficult to achieve. The active ingredients of the rehabilitation process have not been identified either individually or in combination. The placebo condition is therefore, as yet, undefined. Furthermore, from a practical perspective, patients would have to be admitted to a “rehabilitation style” setting for approximately three weeks, where they would participate in activities designed to have no effect on disability, handicap, HRQOL or emotional well-being, for three to four hours per day. In addition, both the patients and staff would need to be unaware that this was the sham intervention. Economic and practical considerations are likely to prevent this from being undertaken in clinical trials of MS inpatient rehabilitation.

This group of MS patients with complex disabilities was chosen since it is these people who are frequently referred for intensive inpatient rehabilitation, and unlike relapsing remitting MS, have limited potential for natural recovery. This sample was similar, in terms of the severity of disease and types of disability, to those of several other rehabilitation studies (e.g. Francabandera et al 1988; Brosseau and Wolfson 1994; Kidd et al 1995; Grasso et al 1996; Fuller et al 1996), suggesting it is representative of the types of patients undertaking rehabilitation. It is appreciated, however, that the generalisability of these results is limited since the study was undertaken in a single specialist centre and all subjects were in the progressive phase and were moderately to severely disabled. There is no assurance that the results will apply to groups free of these selection biases (Nelson et al 1988).
The distinction between the “efficacy” (in research settings) and “effectiveness” (in routine practice) of an intervention reflects the concern that the level and quality of care might be different in research settings than in actual practice (Eddy 1990; D’Agostino et al 1995). In its favour, this study investigated the effectiveness of treatment as implemented in routine clinical practice, rather than a standardised form of intervention designed specifically for the experiment. In some senses this is a weakness because the specific treatment protocols followed are unable to be defined (Pollock et al 1993), but it is also a strength in that this study evaluates “a typical multidisciplinary rehabilitation service” where programmes are tailored to the ever-changing needs of the individual, and standardised prescriptive interventions are actively discouraged.

Previous outcome studies have been criticised for failing to adequately describe the service being evaluated (LaRocca and Kalb 1992). This makes it difficult to pinpoint the features of the programme to which observed effects are attributable, and to replicate the programme elsewhere (Pollock et al 1993). It remains a common problem in rehabilitation research (Wade et al 1992; Gladman et al 1996). Attempts to address this criticism have been made by providing details about the philosophy of the service, the interventions undertaken, the staffing levels, and aspects of service delivery.

Finally, this study demonstrates the effectiveness of rehabilitation in the short term. It is of equal importance to investigate how well the benefits gained are maintained in the longer term. For this purpose a longitudinal study was
undertaken, in parallel with this randomised controlled trial, to observe what happens to these changes in the year following discharge from the rehabilitation unit into the community. The methodology and results of this longitudinal trial are described in the following chapter.

**ii) Outcome measures**

Meaningful interpretation of study results depends largely on the appropriate choice of measurement instrument. Ideally only instruments which have been comprehensively evaluated in terms of their scientific and clinical properties should be used (McDowell and Newell 1987). The current state of the art of measurement in rehabilitation means that this is often not the case. The quality of measurement technology has been generally weak (Turner 1990) and few of the many instruments available for measuring neurological conditions have undergone comprehensive evaluation (Hobart et al 1996b). This makes choosing outcome measures to demonstrate the efficacy of rehabilitation difficult. Compromises are often necessary. A comprehensive understanding of the instruments' strengths and limitations enables aspects of the study design to be altered to limit the impact of these limitations, and interpretation of results to be undertaken in an informed manner.

A number of problems exist with the psychometric properties of the EDSS and FS (refer to Section 2.3.3). Limited inter- and intrarater reliability has been reported (Amato et al 1988; Noseworthy et al 1990; Goodkin et al 1992); validity of the EDSS as an impairment measure has been questioned (Willoughby and Paty 1988); and responsiveness has shown to be low (Hobart
et al 1996c) particularly in the middle ranges of the scale (Willoughby and Paty 1988; Noseworthy et al 1990) where most of this population scored. The impairment scores must therefore be interpreted with some caution. Despite their imperfections, the FS and EDSS are currently the accepted "gold-standard" for measuring disease progression in MS (Hohol et al 1995; Sharrack and Hughes 1996), and no better alternative is available (Whitaker et al 1995; Rudick et al 1996).

To limit problems with the FS and EDSS specific recommendations were adhered to regarding the assessment and interpretation of data (Amato et al 1988; Noseworthy et al 1989, 1990; Verdier-Taillefer et al 1991; Goodkin et al 1992; Whitaker et al 1995). To maximise reproducibility a single neurologist, whose level of intrarater reliability was high (Hobart et al 1996c), rated the patients at each assessment. A shift of one EDSS point and two FS points was required to determine change (Amato et al 1988; Noseworthy et al 1989, 1990; Verdier-Taillefer et al 1991). The FS scales, which are widely accepted as measuring impairment (e.g. Slater 1984; Gulick et al 1993), were used together with the EDSS to provide complimentary information on impairment. The problem of low responsiveness could not be addressed. This means that changes in neurological status may have occurred undetected. The narrow 95% confidence intervals suggest, however, that any such changes, in either direction, were small for both groups.

It is argued, that in the context of clinical trials, responsiveness is the most important measurement property of an instrument (Wilkin et al 1992; Fitpatrick
1994), and may have a considerable impact on the pattern of results obtained
(Fitzpatrick et al 1992a). This is particularly relevant in the evaluation of
interventions for chronic and severely disabled patients, where only marginal
but clinically significant changes may occur (Bindman et al 1990; Hill and
Harries 1994). Low responsiveness may mean that an intervention that
improves health status may show no apparent difference between treated and
untreated patients (type II error).

The SF-36 is considered the gold-standard generic measure of HRQOL. Its
responsiveness however has not been comprehensively evaluated in many
patient groups, including MS. In this study substantial floor effects were
demonstrated on the physical functioning and role limitations dimensions,
suggesting that the SF-36 was unlikely to be responsive to the magnitude of
changes expected in this population. The interpretation of change scores is
therefore limited, and conclusions regarding the effectiveness of rehabilitation
in impacting upon HRQOL should be made with caution. Nevertheless, useful
information has been provided for future researchers with regard to the
relevance and usefulness of the SF-36 in this moderate to severely disabled
population (Freeman et al 1996b). New and better instruments can only be
developed from the proved shortcomings of existing instruments (Hobart et al
1996b).

**iii) Interpretation of changes in outcome scores**

It is important for clinicians to understand changes or differences in health
status scores so that they can interpret research, clinical changes among their
patients, and sample size estimates in clinical trials (Deyo and Patrick 1995). Little attention, however, has been devoted to the interpretation of scores, in a clinical context, in many outcome measures. No evidence exists to determine whether the clinical importance of a given change varies by clinical condition, by different points in the natural history of the disease, or at different levels on the range of the scale. Consequently little is known about what an “x” point change on these measures means from either a clinician’s or patient’s viewpoint. Research is urgently needed to enhance the interpretability of scores in a clinically meaningful context.

Statistical significance does not assure clinical significance (Borenstein 1994). The effect size statistic was therefore used to facilitate interpretation of the results. Although the effect size is useful in supplementing information gained from the statistical significance testing, it is limited in its approach (Kazis et al 1989). It does not provide insight to the meaning of change from either the clinician’s or patient’s viewpoint, and offers little understanding of how a patient appears before and after a certain magnitude of effect. Furthermore, its interpretation is often unfamiliar to many clinicians (Deyo and Patrick 1995). Nevertheless it allowed the change scores to be placed in some perspective by providing an indication of the magnitude of change in relation to the measurement variability.
4.7 Conclusion

This is the first controlled study to investigate the effectiveness of inpatient rehabilitation in MS. Its results have demonstrated that inpatient rehabilitation is effective at reducing disability and handicap in patients with progressive MS, despite unchanging levels of neurological impairment. Levels of health-related quality of life and emotional well-being do not appear to be significantly altered by the rehabilitation process, as measured by the SF-36 and GHQ.

The potential severity of impairment in conjunction with its unpredictable clinical course have lead some health care professionals to view people with MS as poor candidates for rehabilitation (Greenspun et al 1987; BSRM Working Party Report on MS 1993). These results show that patients in the progressive phase of MS, who are identified by an experienced multidisciplinary team as having the potential to benefit from inpatient rehabilitation, are suitable candidates. They demonstrate that even patients with long-standing impairment can make significant functional gains following inpatient rehabilitation.
5.1 Aim of the Study

The primary intentions of this longitudinal study are twofold:

(1) to observe and measure how long any changes made in disability, handicap, health-related quality of life, and emotional well-being in patients with MS are maintained following discharge from inpatient rehabilitation into the community;
(2) to observe the trends in the levels of these outcomes over the twelve month study period.

5.2 Study Design

Single group prospective longitudinal case series.

5.3 Subjects and Methods

Ethical approval was obtained from the Ethics Committee of the National Hospital for Neurology and Neurosurgery in April 1994.

5.3.1 Sample size

The decision to study 50 patients was based on the expected admission rate over the study period (calculated from audit data recording the unit’s levels of patient activity in the previous year).
5.3.2 Recruitment

Subjects were recruited from the weekly multidisciplinary assessment clinic at the National Hospital for Neurology and Neurosurgery (NHNN), Queen Square (refer to section 4.4.2).

i) Inclusion criteria

To exclude as many confounding variables as possible, patients were only entered into the study if they had a diagnosis of clinically or laboratory supported definite MS (Poser et al 1983), were in the progressive phase of the disease as established by the neurologist (Weinshenker et al 1989; Lublin and Reingold 1996), and were considered appropriate for inpatient rehabilitation (Freeman et al 1996a; Johnson and Thompson 1996).

ii) Exclusion criteria

Patients were excluded if they were within one month of relapse, within one month of receiving steroids, had other diseases which may have interfered with outcome, or were cognitively impaired such that they were unable to give informed consent.

5.3.3 Procedures

All patients who met the entry criteria were informed of the study by both written and verbal explanation (Appendix 11). If agreeable to participation, written consent was obtained (The NHNN's “Patient Consent Form for Research on Human Volunteers”).
Patients were assessed at six sequential time points: within 24 hours of admission to the rehabilitation unit, within 24 hours of discharge, and then at three monthly intervals for one year following discharge into the community. Twenty six patients (52%) also participated in the randomised controlled study (Chapter 4). These patients were assessed on one additional occasion, either six weeks prior to or six weeks following rehabilitation, depending upon their group allocation in the randomised controlled study.

5.3.4 Test battery and methods of administration

i) Demographic and diagnostic information

The following information was collected by patient interview and verified by review of the medical notes:

a. Routine demographic and diagnostic data (Appendix 8)

b. Number of relapses in the six months preceding and during the study period (as defined by Poser et al 1983)

c. Pharmacological details during the study period (as detailed in section 4.4.5)

d. Social service and out-patient therapy details (Appendices 12 and 13)

e. Re-hospitalisation and significant changes in social circumstances during the study period (e.g. moved house, divorced).
ii) Measures

The battery of measures and their method of administration was identical to that used in the randomised controlled trial (refer to section 4.4.5). In summary, the measures were:

1. Functional Systems (FS) and Expanded Disability Status Scale (EDSS)
2. The Functional Independence Measure (FIM)
3. The London Handicap Scale (LHS)
4. The Short Form 36 Health Survey Questionnaire (SF-36)
5. The 28-item General Health Questionnaire (GHQ).

At each assessment the measures were administered in a randomised order (from random numbers table, Armitage and Berry 1987) to prevent any systematic bias resulting from practice effects or fatigue (Daly et al 1992). To minimise bias, neither the assessors nor the patients were permitted to see their responses from previous occasions.

5.3.5 The rehabilitation service and programme

All patients participated in the multidisciplinary inpatient rehabilitation programme routinely practised at the NRU, NHNN (refer to section 4.1.6).

5.3.6 Quality control of data input

To ensure accuracy of data input, a detailed visual examination of all data was routinely undertaken, and a random selection of 10% of the data was double checked.
5.3.7 Statistical Analysis

Analysis was undertaken using SPSS Version 6.0 (SPSS Inc.1993).

Descriptive statistics were used to describe the study population in terms of the demographic and disease characteristics, and the performance levels of impairment (EDSS and FS), disability (FIM), handicap (LHS), health-related quality of life (SF-36) and emotional well-being (GHQ). Data was routinely examined for distribution and homogeneity of variance.

Repeated measurements provide an impression of sensitivity in serial design, as well as how closely coupled the measured longitudinal changes are. There is considerable debate about the most appropriate method for analysing the results of repeated measurement studies (Rowell and Walters 1976; Healy 1980; Fleiss 1986; Senn 1994). Several methods have been proposed including: analysis of variance (Aitken 1981); multivariate analysis of variance (Fleiss 1986); latent transition analysis (Collins and Johnston 1995); and the use of summary measures (Matthews et al 1990; Mathews 1993). Analysis of variance, multivariate analysis and linear trend analysis require that the observations are independent and are normally distributed with constant variance (Fleiss 1986; Armitage and Berry 1991). The data did not satisfy these requirements: data at each assessment was not independent since each measurement had been derived from the same patient at each successive time point. The approach was therefore to use summary measures.
Summary measures

Summary measures aim to provide information relevant to the clinical question. Unlike group statistics they consider the individual as the basic unit, relating each individual to their own baseline. In this method the serial measurements from each individual subject are used to construct a single number which summarises the salient aspect of the subject’s response curve (Mathews et al 1990). There are two stages to analysis using summary measures. In the first stage the data points are plotted for every individual to indicate the nature of the relationship with time. From these graphs a summary measure is identified which is appropriate to the clinical question posed. The summary measure is then calculated for each individual. In the second stage the summary measures are analysed.

When considering what summary measure to use in this study a basic assumption was made that patients in the progressive phase of MS would deteriorate over time in a range of daily life outcomes. The question posed was whether the deterioration in disability, handicap, emotional well-being and HRQOL could be delayed following a short period of inpatient rehabilitation. The most clinically relevant summary measure to address this question was the length of time taken (days) for each individual to return to their initial baseline score following discharge. This was the primary criterion for determining the long term benefits of the rehabilitation programme. For every individual, scores of each outcome measure were plotted at the six assessment points. Calculation of the summary measure was determined by measuring the distance along the horizontal axis from discharge to the point at which the line
graph dissected the initial baseline score. The distance in millimetres (mm), where one mm represented two days, was the summary measure for each outcome variable.

Precise calculation of the summary measure proved impossible when all scores remained above the initial baseline value throughout the twelve month study period (FIM n = 13, LHS n = 15, GHQ n = 15, SF36 physical component n = 15, SF36 mental component n = 16). An attempt to overcome this problem was made by modelling each individual’s time series data. By using a Curve Estimation Procedure (SPSS Inc., p 367 - 376) it was anticipated that future readings could be predicted for each individual, thereby providing an estimation of the chosen summary measure. This proved unsuccessful. None of the 11 regression models adequately fitted the data (p > 0.05). At certain time intervals the model always underestimated the scores, while for other intervals the scores were consistently overestimated. This is likely to have been for two main reasons: (1) the relatively small series of observations which made it impossible to extrapolate data points accurately; (2) the large variability in the pattern of each individual’s data throughout the twelve month period. The solution in these cases was to calculate the summary measure as 365 days (the length of the study period), thus providing an estimate of the least number of days that deterioration was delayed. This represents a conservative estimate as it only takes account of the changes made during the year of the study period and excludes any improvements maintained after this time. Following calculation of the summary measure scores for each variable, analysis was undertaken using descriptive statistics.
It is also of interest to know whether patients improved, deteriorated or remained the same compared to their original level of performance at the 12 month assessment. To determine this one year change scores were computed by subtracting baseline scores from the 12 month follow-up assessment scores (this calculation was reversed in computing EDSS and GHQ change scores where higher scores indicate deterioration). Change scores were categorised as “better”, “stayed the same”, or “worse” by comparing them with two standard errors of measurement (SEM), which is approximately the 95% confidence interval for an individual score. For each outcome measure the SEM was calculated according to the formula: \[
SEM = sd \times \sqrt{1 - \text{reliability coefficient}}
\]
where the reliability coefficient was equal to the cronbach’s alpha value reported in the literature (Ware 1994a). This method enables some of the “noise” associated with repeated assessments to be taken into account by considering the fluctuations due to measurement error (Ware 1994b). It thereby increases the confidence with which individual change scores can be interpreted. According to this method, patients were classified as “worse” if their change score indicated a decline greater than two SEM’s. They were classified as “stayed the same” if their change score was within two SEM’s, and as “better” if their change score was greater then two SEM’s.

Correlation analyses were used to analyse the relationships between: (1) each of the outcomes measured on admission; (2) admission scores for all outcome measures and the magnitude of benefits gained during the inpatient stay; (3) the baseline level of performance and the length of time benefits were carried over; (4) the magnitude of benefits gained during the inpatient stay and the
length of time these benefits were maintained in the longer term. Pearson's product moment correlation coefficient was used with interval data and Spearman's rank correlation coefficient was used with ordinal data. The strength of the relationships were defined according to McHorney and colleagues (1993) where the relationship is strong if the correlation is greater than 0.70, moderate to substantial if the correlation is 0.30 to 0.70, and weak if the correlation is less than 0.30.

All available data was analysed at each time point: (admission and discharge n = 50; three month n = 48; six, nine and 12 month n = 46).

5.4 Results

Eligible patients were admitted sequentially into this study until 50 patients, who fulfilled the entry criteria, were recruited. The recruitment period was from May 1994 - May 1995. During this time 186 consecutive MS patients were assessed, of whom 73 patients were selected for admission for inpatient rehabilitation.

Twenty two patients did not satisfy the entry criteria (Table 5.1) and one person declined to participate in the study.

Table 5.1 Patients failing to meet the study's entry criteria (n = 22)

<table>
<thead>
<tr>
<th>Reasons for failing to meet criteria</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>On steroids / in acute relapse</td>
<td>14</td>
</tr>
<tr>
<td>Within one month of relapse</td>
<td>2</td>
</tr>
<tr>
<td>Not definite MS</td>
<td>2</td>
</tr>
<tr>
<td>Severe cognitive disturbance</td>
<td>2</td>
</tr>
<tr>
<td>Other diseases which may interfere with outcome (peripheral demyelination x1; poorly controlled epilepsy x1)</td>
<td>2</td>
</tr>
</tbody>
</table>
Of the 50 patients entering the study, complete data for every assessment point was available on 44 patients. Details of the six patients with incomplete data are summarised below:

- **Case 4** was a 56 year old male with primary progressive MS, in the advanced stages of the disease (admission scores: EDSS 8.0; FIM 36; LHS 65.7; GHQ 8). Three months following discharge from the unit Mr N was re-admitted to hospital for six months. He died of pneumonia prior to the final assessment.

- **Case 9** was a 36 year old female with secondary progressive MS, who was severely disabled (admission scores: EDSS 9.0; FIM 20; LHS 31.6; GHQ 23). Following the three month assessment Ms M withdrew from the study, finding the questions too distressing.

- **Case 19** was a 54 year old male with secondary progressive MS, in the advanced stages of the disease (admission scores: EDSS 8.0; FIM 20; LHS 38.1; GHQ 5). Prior to admission he had been bed bound at home for eight months. Mr C moved from London to an adapted bungalow in Gloucester one month following discharge. He declined the six and nine month assessments, finding the burden of attending the outpatient clinic too great. The 12 month assessment was undertaken at his home in Gloucester.

- **Case 24** was a 37 year old male with secondary progressive MS, in the advanced stages of the disease (admission scores: EDSS 8.5; FIM 14; LHS 67.9; GHQ 19). Six months following discharge he suffered a severe relapse and was transferred into a nursing home for long term care. He died prior to the 12 month assessment.

- **Case 26** was a 43 year old female with secondary progressive MS, in the advanced stages of the disease (admission scores: EDSS 9.0; FIM 13; LHS
62.2; GHQ 17). She declined the three, six and nine month assessments, finding the burden of attending the outpatient clinic too great. The 12 month assessment was undertaken at her home in Surrey.

- Case 38 was a 57 year old male with primary progressive MS (admission scores: EDSS 6.5; FIM 62; LHS 58.6; GHQ 9). He was unable to be traced following discharge despite repeated attempts via post and telephone.

In summary, of the six patients with incomplete data all except one was severely disabled. Two died shortly before the final assessment, three declined a proportion of the reviews, and one was lost to follow-up.

5.4.1 Sample baseline characteristics

i) Demographic and diagnostic characteristics

On admission, the mean age of the subjects was 44.8 years (sd +/-9.7, range 25 - 66), and 29 (58%) were female. The disease characteristics are summarised in Table 5.2.

Thirty seven patients were married, seven were single, and six were either separated or divorced. Thirty two patients were retired (31 of them on medical grounds), eight were employed, and four were full-time homemakers. Among the six unemployed patients four were not currently seeking work.
Table 5.2 Baseline disease characteristics

<table>
<thead>
<tr>
<th>Variable</th>
<th>n = 50</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disease pattern: secondary progressive</td>
<td>86%</td>
</tr>
<tr>
<td>primary progressive</td>
<td>14%</td>
</tr>
<tr>
<td>Mean years since first symptoms: +/- sd, (range)</td>
<td>15.3, +/-8.8, (3 - 36)</td>
</tr>
<tr>
<td>Mean years since diagnosis: +/-sd, (range)</td>
<td>11.6, +/-8.2, (1.5 - 35)</td>
</tr>
<tr>
<td>Number of patients reporting a relapse in the 6 months preceding initial assessment</td>
<td>12</td>
</tr>
<tr>
<td>Mean number of months since relapse (if in last 6 months): +/-sd (range)</td>
<td>4.1, +/-1.9 , (1.5 - 6)</td>
</tr>
<tr>
<td>Progression of symptoms in past 6 months</td>
<td>74%</td>
</tr>
</tbody>
</table>

Thirty two patients lived in their own home, 14 in rented accommodation, three in their parents home, and one in long term residential care. In total 43 patients lived with a spouse or family member, six lived alone, and one in a residential unit. Forty five patients required some assistance with their self care. In 19 cases this was provided by a family member or friend, and in 26 cases additional care was required by social services or privately paid help. The specific hospital and community social services received at the time of admission are outlined in Table 5.3.
Table 5.3 Hospital and Community services on admission

<table>
<thead>
<tr>
<th>Services</th>
<th>% patients (n = 50)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Community occupational therapy</td>
<td>40</td>
</tr>
<tr>
<td>Home help</td>
<td>38</td>
</tr>
<tr>
<td>District nurse</td>
<td>30</td>
</tr>
<tr>
<td>Community physiotherapy</td>
<td>20</td>
</tr>
<tr>
<td>Social worker</td>
<td>16</td>
</tr>
<tr>
<td>Day care</td>
<td>14</td>
</tr>
<tr>
<td>Wheelchair services clinic</td>
<td>10</td>
</tr>
<tr>
<td>Other (cognitive therapist, care manager, independent living team)</td>
<td>8</td>
</tr>
<tr>
<td>Regular outpatient physiotherapy</td>
<td>8</td>
</tr>
<tr>
<td>Personal care attendant</td>
<td>6</td>
</tr>
<tr>
<td>Outpatient physiotherapy review</td>
<td>4</td>
</tr>
<tr>
<td>Day care</td>
<td>2</td>
</tr>
<tr>
<td>Speech therapy</td>
<td>0</td>
</tr>
</tbody>
</table>

Of the 50 patients, 35 were admitted from home, 13 from an acute hospital setting, one from another rehabilitation unit and one from a residential unit. On discharge 47 patients returned home, one to an acute hospital, one to a residential unit, and one to her parents home.

ii) Outcome measure scores

The baseline values for all outcome measures are summarised in Table 5.4.
Table 5.4 Baseline outcome measurement values

<table>
<thead>
<tr>
<th>Outcome measure</th>
<th>Measurement score (n = 50)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Kurtzke’s Scales:</strong></td>
<td></td>
</tr>
<tr>
<td>EDSS median, mode, range (0-10)</td>
<td>6.8 6.5 (6.0 - 9.0)</td>
</tr>
<tr>
<td><strong>Functional Systems:</strong></td>
<td></td>
</tr>
<tr>
<td>pyramidal (0-6)</td>
<td>4 4 (1 - 5)</td>
</tr>
<tr>
<td>cerebellar (0-5)</td>
<td>2 3 (0 - 5)</td>
</tr>
<tr>
<td>brainstem (0-5)</td>
<td>2 1 (0 - 4)</td>
</tr>
<tr>
<td>sensory (0-6)</td>
<td>2 3 (0 - 5)</td>
</tr>
<tr>
<td>sphincter (0-6)</td>
<td>2 2 (0 - 6)</td>
</tr>
<tr>
<td>visual (0-6)</td>
<td>2 0 (0 - 6)</td>
</tr>
<tr>
<td>mental (0-5)</td>
<td>1 0 (0 - 4)</td>
</tr>
<tr>
<td><strong>FIM (motor domain)</strong></td>
<td>61.5 36.8 - 75 (13 - 87)</td>
</tr>
<tr>
<td><strong>LHS (total score)</strong></td>
<td>60.3 +/-13.1 (31.6 - 87.2)</td>
</tr>
<tr>
<td><strong>SF-36 domains:</strong></td>
<td></td>
</tr>
<tr>
<td>physical function (0-100)</td>
<td>9.5 +/-12.7 (0 - 70)</td>
</tr>
<tr>
<td>mental health (0-100)</td>
<td>57.5 +/-20.4 (4 - 92)</td>
</tr>
<tr>
<td>role limitation-physical (0-100)</td>
<td>15.0 +/-29.5 (0 - 100)</td>
</tr>
<tr>
<td>role limitation- mental (0-100)</td>
<td>39.3 +/-41.9 (0 - 100)</td>
</tr>
<tr>
<td>energy and vitality (0-100)</td>
<td>35.1 +/-21.2 (0 - 75)</td>
</tr>
<tr>
<td>social functioning (0-100)</td>
<td>39.3 +/-25.5 (0 - 89)</td>
</tr>
<tr>
<td>pain (0-100)</td>
<td>65.6 +/-29.0 (0 - 100)</td>
</tr>
<tr>
<td>health perceptions (0-100)</td>
<td>36.4 +/-20.2 (0 - 100)</td>
</tr>
<tr>
<td>physical health component summary score (0-100)</td>
<td>27.8 +/-6.9 (12.2 - 43.7)</td>
</tr>
<tr>
<td>mental health component summary score (0-100)</td>
<td>39.2 +/-9.6 (14.5 - 63.6)</td>
</tr>
<tr>
<td><strong>GHQ:</strong> median, IQR, range (0-28)</td>
<td>9.5 3.8 - 17 (0 - 28)</td>
</tr>
</tbody>
</table>

IQR = interquartile range; SF-36 scores are transformed scores
The severity of disease is demonstrated by the generally high FS and EDSS scores. In pyramidal functions 66% of the cohort had a marked paraparesis, paraplegia, hemiplegia, or quadriplegia (scores 4-6). In cerebellar functions 48% had moderate or worse truncal or limb ataxia or severe ataxia of all limbs (scores 3-5). In brainstem functions 24% showed severe brain-stem abnormalities including extra-ocular motor weakness, dysarthria or dysphagia (scores 3-5). In sensory functions 46% showed severe loss of vibration, proprioception, or pain sensation (scores 3-6). In bowel and bladder functions 48% had frequent urinary incontinence, need for catheterisation, or need for almost constant use of measures to evacuate bowels (scores 3-6). In visual functions 12% had corrected visual acuity worse than 20/100 in either eye (scores 4-6). In mental functions one patient had a marked decrease in mentation (scores 4 and 5). The high EDSS scores show that in terms of locomotion, at best, patients required intermittent or unilateral constant assistance to walk 100 metres, while at worst they were essentially bed bound. It is notable that 23 patients (46%) scored ≥ 7.5 points on the EDSS, indicating that they were essentially restricted to the wheelchair.

The FIM scores were distributed across virtually the entire spectrum of the scale (13-91), demonstrating the wide ranging levels of dependence for activities of daily living. Physical assistance from another person (scores 1-4) was required by: 68% of patients with bath transfers; 58% with bed transfers; 56% in dressing; 54% in bladder management; 48% in bathing; 40% with toilet transfers; 36% in toileting; 28% in feeding and grooming; 20% in bowel management. In terms of
locomotion 46% relied upon a wheelchair for mobility indoors. Fifty six percent of patients were unable to climb stairs, despite maximal physical assistance.

The distribution of the LHS scores was restricted to the middle portion of the range. At worst, patients scored a total of 31.6 points, and at best 87.2 points.

The profile of scores indicate the areas where this group was most handicapped. In mobility functions 72% of the population, at best, were limited “quite a lot” in their ability to get of the house (scores 3-6). In physical independence 52% required help in looking after themselves more than once a day (scores 4-6). In occupation 44% were “unable to do a lot of things but found something to do most of the time” (score 4), while 24% were “only able to find something to do some of the time” (score 5). Social life was “limited slightly by health” in 50% of patients (score 2), and to the extent that “they felt uncomfortable with strangers” in 24% of cases (score 3). Problems with hearing, speaking, seeing or memory were reported by 48% (scores 2-6), although this was “only slight” in 30% of patients (score 2). With regard to economic considerations, 26% of patients relied on state benefits for financial support.

As might be anticipated (Brunet et al 1996; Freeman et al 1996b; Hermann et al 1996), the SF-36 profile of scores demonstrated particularly severe problems with physical functioning and role limitations due to physical and mental problems. A marked floor effect was observed in these three dimensions where the lowest possible score of zero was scored by 72% of patients for role limitations due to physical problems, 46% for role limitations due to mental problems, and 36% for physical functioning. Moderately severe problems were
demonstrated in the dimensions of energy and vitality, social functioning, and health perceptions.

The GHQ scores confirm the presence of problems in the area of emotional health and well-being. Using the criterion established by Goldberg (1978, refer to Section 4.5.1), 74% of patients scored equal to or greater than five points on the total GHQ score, indicating the presence of emotional disturbance. Of these, 34% scored equal to or greater than 15 points indicating a “markedly abnormal” level of emotional disturbance. Vickrey and colleagues (1995) used a “cut-off” score of 66 on the mental health scale of the SF-36 to divide their population into those who had scores which were associated with depressive symptoms and those who did not. Using this criterion, 72% of patients demonstrated depressive symptoms.

5.4.2 The rehabilitation intervention

i) Rehabilitation programme

All patients received an individualised, goal-oriented programme (as detailed in Section 4.4.6) addressing a wide range of areas (Table 5.5) for an average of 23 days (median 19, sd +/-11.5, range 10 - 62). All patients had medical, nursing, occupational therapy and physiotherapy input. Seventy percent of patients were assessed by the neuropsychologist, 66% by the social worker, and 30% percent by the speech and language therapist. Consultation was obtained as required from psychiatric, neuro-urological and dietetic services.
Table 5.5 Broad areas addressed in the rehabilitation programme

<table>
<thead>
<tr>
<th>Broad areas addressed</th>
<th>% patients (n = 50)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Re-education posture and balance</td>
<td>86</td>
</tr>
<tr>
<td>Personal activities of daily living</td>
<td>82</td>
</tr>
<tr>
<td>Provision of home exercise programme</td>
<td>76</td>
</tr>
<tr>
<td>Posture and seating</td>
<td>74</td>
</tr>
<tr>
<td>Cognitive assessment</td>
<td>72</td>
</tr>
<tr>
<td>Bladder management</td>
<td>72</td>
</tr>
<tr>
<td>Fatigue management</td>
<td>72</td>
</tr>
<tr>
<td>Leisure pursuits</td>
<td>72</td>
</tr>
<tr>
<td>Emotional assessment and counselling</td>
<td>66</td>
</tr>
<tr>
<td>Advice to carers</td>
<td>62</td>
</tr>
<tr>
<td>Domestic activities of daily living</td>
<td>58</td>
</tr>
<tr>
<td>Gait re-education</td>
<td>56</td>
</tr>
<tr>
<td>Drug management</td>
<td>56</td>
</tr>
<tr>
<td>Bowel management</td>
<td>54</td>
</tr>
<tr>
<td>Tone management</td>
<td>52</td>
</tr>
<tr>
<td>Wheelchair skills</td>
<td>52</td>
</tr>
<tr>
<td>Management of contractures</td>
<td>32</td>
</tr>
<tr>
<td>Work skills</td>
<td>30</td>
</tr>
<tr>
<td>Communication skills</td>
<td>22</td>
</tr>
<tr>
<td>Swallowing assessment and management</td>
<td>22</td>
</tr>
</tbody>
</table>

ii) Recommendations on discharge

An important goal of the rehabilitation team may be to set up a "network of community support" to maintain the person in their home environment. This support is often especially necessary when the disease is in the advanced stages. Additional hospital and community services recommended by the multidisciplinary team following discharge, and the length of time taken for them to be commenced, were therefore recorded (Table 5.6)
### Table 5.6: Additional hospital and community services recommended at discharge

<table>
<thead>
<tr>
<th>Services</th>
<th>Patients (n = 50)</th>
<th>No. of patients who had received initial contact by the service in each 3 month period:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>3mo.</td>
</tr>
<tr>
<td>Wheelchair services clinic</td>
<td>27</td>
<td>21</td>
</tr>
<tr>
<td>Community occupational therapy</td>
<td>21</td>
<td>19</td>
</tr>
<tr>
<td>District nurse</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>Regular outpatient physiotherapy</td>
<td>12</td>
<td>8</td>
</tr>
<tr>
<td>Other (cognitive therapist, pain specialist, environmental controls, dietician, bowel surgery, orthotics etc.)</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>Outpatient physiotherapy review</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>Community physiotherapy</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>Psychology / Counselling</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>Social worker</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>Home help</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Care manager</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>Day care</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Re-housing</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Respite care</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

mo. = months

In addition, a range of house adaptations was recommended. These included: rails (n=12), level access shower (n=4), stairlift (n=3), ramp (n=5), fixed overhead hoist (n=2), step widening (n=1), and door entry system (n=3). In 20 cases these adaptations were commenced by three months, in five cases by six months, in one case by nine months, and in two cases by 12 months following discharge from the unit. In two cases the adaptations had not been commenced within the study period.
### iii) Drug management

On initial assessment 98% of patients were taking some form of prescribed medication (Table 5.7). The most common changes in medication during the study period were in bowel management, anti-depressant medication and analgesics. It is notable that over the 12 months: five patients had a three day course of intravenous methyl prednisolone; four patients commenced Cari Loders Drug Regime (70 m.g. lofepramine, 500 m.g. α phenylalanine B.D., 1000 m.g. B12 injection once/week), of which two had stopped it within three months; two patients were admitted into the interferon Beta-1b placebo controlled trial at six months; one person underwent a phenol block to manage severe lower limb flexor spasms.

<table>
<thead>
<tr>
<th>Drug category</th>
<th>adm.</th>
<th>d/c</th>
<th>3 mo.</th>
<th>6 mo.</th>
<th>9 mo.</th>
<th>12 mo.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Immunosuppressants</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Steroids</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Anti-spasticity agents</td>
<td>44</td>
<td>48</td>
<td>52</td>
<td>54</td>
<td>56</td>
<td>52</td>
</tr>
<tr>
<td>Anti-depressants</td>
<td>18</td>
<td>28</td>
<td>29</td>
<td>35</td>
<td>37</td>
<td>35</td>
</tr>
<tr>
<td>Anti-tremor agents</td>
<td>10</td>
<td>12</td>
<td>10</td>
<td>11</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>Fatigue</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Bowel management</td>
<td>40</td>
<td>60</td>
<td>54</td>
<td>56</td>
<td>61</td>
<td>63</td>
</tr>
<tr>
<td>Bladder management</td>
<td>34</td>
<td>36</td>
<td>38</td>
<td>39</td>
<td>41</td>
<td>41</td>
</tr>
<tr>
<td>Analgesics/Sensory disturbances</td>
<td>22</td>
<td>32</td>
<td>35</td>
<td>37</td>
<td>48</td>
<td>40</td>
</tr>
<tr>
<td>Other</td>
<td>42</td>
<td>50</td>
<td>58</td>
<td>48</td>
<td>46</td>
<td>44</td>
</tr>
</tbody>
</table>

adm. = admission; d/c = discharge; mo. = months
5.4.3 Self-reported relapses

Of the 46 patients with complete data (including the two who died), fifteen patients reported a total of 22 relapses over the study period. Of these, 10 patients suffered one relapse, three patients suffered two relapses, and two patients suffered three relapses.

5.4.4 Re-hospitalisations

Table 5.8 details the number of times patients were re-hospitalised over the study period, and the reasons why this occurred. In addition one person was admitted to a residential unit for respite care for two weeks every two months.

<table>
<thead>
<tr>
<th>Reasons for re-hospitalisation</th>
<th>No. of patients (n = 46)</th>
<th>Length of stay (mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intravenous methyl prednisolone</td>
<td>2</td>
<td>3 days</td>
</tr>
<tr>
<td>Cellulitis</td>
<td>1</td>
<td>10 days</td>
</tr>
<tr>
<td>Assessment for thalamotomy (surgery not undertaken)</td>
<td>1</td>
<td>2 weeks</td>
</tr>
<tr>
<td>Appendectomy</td>
<td>1</td>
<td>2 weeks</td>
</tr>
<tr>
<td>Rehabilitation at local hospital</td>
<td>2</td>
<td>3 weeks</td>
</tr>
<tr>
<td>Surgery for colostomy and suprapubic catheter</td>
<td>1</td>
<td>4 weeks</td>
</tr>
<tr>
<td>Severe relapse</td>
<td>6</td>
<td>7.5 weeks (range 2-18)</td>
</tr>
<tr>
<td>Uncontrolled epilepsy</td>
<td>1</td>
<td>8 weeks</td>
</tr>
<tr>
<td>Control of severe pain and spasms</td>
<td>1</td>
<td>6 months</td>
</tr>
</tbody>
</table>

5.4.5 Outcome and Summary Measure scores

The outcome measurement scores define the level in each of the dimensions measured. The summary measure scores define the carryover of benefit following discharge into the community. They determine the time taken (days)
for an individuals outcome measure score to return to the baseline level, thus reflecting the delayed deterioration associated with inpatient rehabilitation. Outcome measure and summary measure scores were calculated for impairment, disability, handicap, HRQOL and emotional well-being.

i) Neurological status

The median EDSS scores at each assessment were plotted for the group (Figure 5.1). This graph demonstrates a gradual deterioration in the groups overall level of neurological status during the study period.

**Figure 5.1** Group median EDSS scores at each assessment point

The change in distribution of the EDSS scores over the study period (Figure 5.2) confirms the shift towards higher scores at 12 months compared to admission.
Figure 5.2 Distribution of the group's EDSS scores on admission and 12 months

The overall impression of a decline in neurological status is further supported by the distribution of FS scores over the study period (Table 5.9). Compared to admission more patients score at higher levels on the FS scales at 12 months.

The variability of the disease course is a key feature of MS. Hence, in addition to the group scores, graphs of each individual's EDSS scores were plotted in relation to time (Figure 5.3). The graphs show that 29% of patients increased their EDSS score by one point or more over the study period, indicating a clinically significant deterioration in neurological status. Although fluctuations occurred during the 12 months in 54% of patients, their final score remained within +/- 0.5 points of the initial value. In 13% of patients the score remained identical at each assessment point (two of whom scored 6.0 and four of whom scored 6.5). In two patients (4%) overall improvement in neurological status was observed at the 12 month assessment as demonstrated by a reduction of one EDSS point or greater.
Table 5.9 Functional Systems scores for the group at each assessment point

<table>
<thead>
<tr>
<th>Functional Systems</th>
<th>adm. (n = 50)</th>
<th>d/c (n = 50)</th>
<th>3 mo. (n = 48)</th>
<th>6 mo. (n = 46)</th>
<th>9 mo. (n = 46)</th>
<th>12 mo. (n = 46)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pyramidal ≥ 4</td>
<td>66</td>
<td>70</td>
<td>75</td>
<td>73</td>
<td>80</td>
<td>76</td>
</tr>
<tr>
<td>cerebellar ≥ 3</td>
<td>48</td>
<td>46</td>
<td>44</td>
<td>48</td>
<td>52</td>
<td>48</td>
</tr>
<tr>
<td>brainstem ≥ 3</td>
<td>24</td>
<td>20</td>
<td>29</td>
<td>27</td>
<td>29</td>
<td>25</td>
</tr>
<tr>
<td>sensory ≥ 3</td>
<td>46</td>
<td>46</td>
<td>48</td>
<td>54</td>
<td>49</td>
<td>58</td>
</tr>
<tr>
<td>sphincter ≥ 3</td>
<td>48</td>
<td>52</td>
<td>60</td>
<td>64</td>
<td>71</td>
<td>60</td>
</tr>
<tr>
<td>visual ≥ 4</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>20</td>
<td>15</td>
<td>10</td>
</tr>
<tr>
<td>mental ≥ 4</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>6</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

adm. = admission; d/c = discharge; mo. = months
Fig 5.3 Individual plots of EDSS scores against time (n = 46)

a/d = admission; d/c = discharge; lower scores indicate improvement; _____ = baseline score reference line
ii) Disability

The median FIM motor domain scores were plotted for the group at each assessment point (Figure 5.4). Initial improvement in overall disability occurred during the rehabilitation period (median admission score 61.5, range 13-87; median discharge score 74, range 13-90). Following discharge the level of performance steadily deteriorated throughout the study period. At nine months the median FIM score was 65.0 (range 13-88), and by 12 months it was only two points above the initial baseline (median 63.5, range 13-90).

**Figure 5.4** Group median FIM motor scores at each assessment point

Each individual’s FIM scores were plotted in relation to time (Figure 5.5). As expected in MS a wide variety of curves was observed including multiple peaks (case 41 and 46), sustained improvement (case 14), and steady deterioration (case 11). In ten cases a marked deterioration in performance (loss of 15 points or more) occurred between consecutive assessment points. This coincided with a relapse in five cases and with an appendectomy in one case.
Fig 5.5 Individual plots of FIM scores against time (n = 46)

 parade = admission; d/c = discharge; higher scores indicate improvement; = baseline score reference line; N.B. Cases 4 & 24 died prior to 12 month assessment, therefore serial data to 9 months graphed
Summary measures were calculated for each individual to determine how long the deterioration in disability was delayed following completion of the rehabilitation programme (mean 186 days, range 0 - 365, sd +/-150). The large range and standard deviations indicate that marked variability in the long term maintenance of benefits occurred between individuals.

One year change scores were computed by subtracting each patient's baseline FIM score from the corresponding 12 month score. The change in performance was then categorised by comparing the magnitude of the change score to two times the size of the SEM (where the SEM for FIM = +/- 4.0). Patients with change scores above +8.0 were classified as "better". Patients with change scores below -8.0 were classified as "worse". Patients with change scores between +/- 8.0 were classified as "stayed the same". According to this criterion, 16% of patients were better, 30% were worse, and 54% were the same at the 12 month assessment compared to their baseline.

It was notable that the change in FIM scores tended to be large in those patients categorised to the "worse" group (mean loss 22.5, range -8 - -50, sd +/-14.1), with ten of the 15 patients (67%) deteriorating by ≥ -15 FIM points. In contrast the gains made by those categorised to the "better" group were more modest (mean gain 14, range +7 - + 41, sd +/-10.5), with only three of the 10 patients (33%) improving by ≥ +15 FIM points.
iii) Handicap

The mean LHS scores were plotted for the group at each assessment point (Figure 5.6). Initial improvement in overall handicap occurred during the inpatient rehabilitation period (mean admission score 60.3, sd +/-13.1, range 31.6-87.2; mean discharge score 64.4, sd +/-11.8, range 26.1-87.2). Following discharge the group’s level of performance steadily declined (mean three month score 63.9, sd +/-14.5, range 16.3-92.6; mean six month score 63.4, sd +/-13.1, range 29.1-89.9). By the nine month assessment the scores had returned to only marginally above the baseline level (mean nine month score 61.2, sd +/-15.7, range 25-90.3; mean 12 month score 61.6, sd +/-13.7, range 32.1-87.5).

**Figure 5.6** Group mean LHS scores at each assessment point

Graphs of each individual’s LHS scores were plotted in relation to time (Figure 5.7).
Fig 5.7 Individual plots of LHS scores against time (n = 46)

ad = admission, d/c = discharge, higher scores indicate improvement; baseline score reference line;
N.B. Cases 4 & 24 died prior to 12 month assessment, therefore serial data to 9 months graphed
The individual graphs show that the majority of the 35 patients who improved their overall level of handicap peaked once they had returned into the community. Nine patients scored highest in handicap at discharge, 12 peaked at three months, five at six months, four at nine months, and five at 12 months.

Summary measures were calculated for each individual to determine how long the deterioration of handicap was delayed following the rehabilitation programme (mean 185 days, range 0 - 365, sd +/-162). As in disability, marked variability in both the pattern of performance and the length of carryover was noted between individuals.

One year change scores were computed by subtracting each patient’s baseline LHS score from the corresponding 12 month score. The change in performance was then categorised by comparing the magnitude of the change score to the size of two times the SEM (where the SEM for LHS = +/-4.5). Patients with change scores above +9.0 were classified as “better”. Patients with change scores below -9.0 were classified as “worse”. Patients with change scores between +/- 9.0 were classified as “stayed the same”. According to this criterion, 22% of patients were better, 13% were worse, and 65% were the same at the 12 month assessment compared to their baseline.

iv) Health-related Quality Of Life

The mean SF-36 physical and mental health summary scores were plotted for the group at each assessment point (Figures 5.8a,b). Graphs of each
individual's SF-36 scores (mental and physical component) were plotted in relation to time (Figure 5.9a,b).

**Figure 5.8 a,b** Group mean SF-36 Physical and Mental Health Summary scores at each assessment point

![Graphs showing SF-36 Physical and Mental Health Summary scores over time with 95% confidence intervals.](image)

N.B. Higher scores imply improvement

Different patterns in the reported levels of HRQOL were noted between the two dimensions. In the physical component over half of the patients (54.4%) reported their HRQOL to be best at three months, with 28.2% reporting this to be the case at six months. In contrast, in the mental component 21% of patients reported their HRQOL to be best at three months, with the majority (61%) reporting it to be best at six months. In both dimensions the group demonstrated a marked reduction in HRQOL at nine months. In the physical component the group’s score fell marginally below the baseline level (mean loss 0.7 points). The group’s mental scores however remained above the baseline level for the entire study period.

Summary measures were calculated for each individual to determine how long the deterioration in HRQOL was delayed following the rehabilitation programme. The average delay was 293 days (sd +/-56, range 190 - 365) for the physical component, and 171 days (sd +/-171, range 0 - 365) for the mental component.
Fig 5.9a Individual plots of SF-36 Physical Health Summary scores against time (n = 46)

T (months)
Mean of group

a/d = admission; d/c = discharge; higher scores indicate improvement; baseline score reference line;
N.B. Cases 4 & 24 died prior to 12 month assessment, therefore serial data to 9 months graphed
Fig 5.9b Individual plots of SF-36 Mental Health Summary scores against time
(n = 46)

Mean of group

a/d = admission, d/c = discharge, higher scores indicate improvement, baseline score reference line;
N.B. Cases 4 & 24 died prior to 12 month assessment, therefore serial data to 9 months graphed
One year change scores were computed by subtracting each patient’s baseline SF-36 health summary score from the corresponding 12 month score. The change in performance was then categorised by comparing the magnitude of the change score to two times the size of the SEM (where the SEM for the physical component = +/-2.8, and for the mental component = +/-3.2). Patients with change scores above +/-5.6 for the physical component and +/-6.4 for the mental component were classified as “better”. Patients with change scores between +/-5.6 for the physical component and +/-6.4 for the mental component were classified as “stayed the same”. Patients with change scores below -5.6 for the physical component and -6.4 points for the mental component were classified as “worse” (Table 5.10).

Table 5.10 SF-36 Physical and Mental Health Summary change scores

<table>
<thead>
<tr>
<th></th>
<th>Better (%)</th>
<th>Stayed the same (%)</th>
<th>Worse (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical Health Summary score</td>
<td>18</td>
<td>61</td>
<td>21</td>
</tr>
<tr>
<td>Mental Health Summary score</td>
<td>48</td>
<td>32</td>
<td>20</td>
</tr>
</tbody>
</table>

v) Emotional well-being

The median GHQ scores were plotted for the group at each assessment point (Figure 5.10). Initial improvement in emotional well-being occurred during the inpatient rehabilitation period (median admission score 9.5, range 0-28; median discharge score 1.5, range 0-24). Although this level deteriorated slightly at three months (median 4.0, range 0-25), it then remained reasonably stable throughout
the rest of the study period (six months median 5, range 0.25; nine months median 4, range 0-27; twelve month median 4, range 0-28). The median group scores on the GHQ remained below the baseline score for the entire 12 months, indicating a sustained improvement in emotional well-being compared to baseline values.

**Figure 5.10** Group median GHQ scores at each assessment point

Graphs of each individual's GHQ scores were plotted in relation to time (Figure 5.11). Summary measures were calculated for each individual to determine how long the deterioration of emotional well-being was delayed following the rehabilitation programme (mean 208 days, sd +/-153, range 0 - 365).
Fig 5.11 Individual plots of GHQ scores against time (n = 46)

a/d = admission, d/c = discharge; lower scores indicate improvement; = baseline score reference line
Cases 4 & 24 died prior to 12 month assessment, therefore serial data to 9 months graphed
One year change scores were computed by subtracting each patient’s 12 month GHQ scores from the corresponding baseline scores. The change in performance was then categorised by comparing the magnitude of the change score to two times the size of the SEM (where the SEM for GHQ = +/-2.6). Patients with change scores above +5.2 were classified as "better". Patients with change scores below -5.2 were classified as "worse". Patients with change scores between +/- 5.2 were classified as "stayed the same". According to this criterion, 33% of patients were better, 9% were worse, and 58% stayed the same at the 12 month assessment.

Using Goldberg’s criterion, the percentage of patients who experienced emotional disturbance (> 5 GHQ points) decreased over the study period. On admission 74% of patients scored greater than five points, indicating the presence of emotional disturbance. This had reduced to 47% of patients at the 12 month assessment. Of these patients, the percentage who were described as experiencing “markedly abnormal” levels of emotional disturbance (≥ 15 GHQ points) also decreased during the study period. At admission 34% of patients were in this category compared to 15% at the 12 month assessment.
Table 5.11 Summary of data for all outcome measures assessed in the longitudinal study

<table>
<thead>
<tr>
<th>Outcome measure</th>
<th>Time to return to baseline score (days)</th>
<th>Categorisation of patients according to change scores +/- 2 SEM's (% of patients)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean +/- sd</td>
<td>range</td>
</tr>
<tr>
<td>EDSS</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>FIM</td>
<td>186 +/-150</td>
<td>0-365</td>
</tr>
<tr>
<td>LHS</td>
<td>185 +/-162</td>
<td>0-365</td>
</tr>
<tr>
<td>SF36-physical</td>
<td>293 +/-56</td>
<td>190-365</td>
</tr>
<tr>
<td>SF36-mental</td>
<td>171 +/-171</td>
<td>0-365</td>
</tr>
<tr>
<td>GHQ</td>
<td>208 +/-153</td>
<td>0-365</td>
</tr>
</tbody>
</table>

SEM = Standard Error of Measurement; Change scores computed by subtracting baseline score from 12 month score
5.4.6 Correlation analyses

i) Association between outcome measures

Associations between each of the measures on admission were determined using Pearson's product moment correlation coefficient for the interval data, and Spearman's rank correlation coefficient for the ordinal data (Table 5.12). A strong correlation was observed between overall severity of disease (EDSS) and disability. Handicap was moderately associated with a number of the measures - overall disease severity, disability, the mental component of HRQOL, and emotional well-being. There were also moderate correlations between the physical component of HRQOL and disability, and the mental component of HRQOL and emotional well-being.

In terms of the specific neurological impairments (as measured by FS scores): brainstem functions correlated moderately with handicap and emotional well-being; sphincter control and pyramidal functions correlated moderately with overall disease severity and disability; sensory functions correlated moderately with overall disease severity and weakly with disability; visual functions correlated moderately with all outcomes other than the physical component of HRQOL.
Table 5.12 Correlation matrix of baseline outcome measurement scores.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>EDSS</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FIM</td>
<td>-0.8271</td>
<td>-0.3042</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LHS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GHQ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF36-physical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF36-mental</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FS: brainstem</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FS: sphincter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FS: cerebell</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FS: pyramidal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FS: sensory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FS: mental</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FS: visual</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Correlation coefficients and p-values provided.
ii) Associations between admission scores and the magnitude of changes made during rehabilitation

These relationships were explored by correlating admission scores on each outcome measure with the change scores (computed by subtracting admission scores from discharge scores for FIM, LHS and SF-36, and the reverse for EDSS and GHQ). The EDSS score did not correlate with changes on any outcome measure. The only FS scale to demonstrate a significant relationship with a change score was the Mental System with the GHQ change scores ($r = -0.3779$, $p = 0.007$).

Moderate correlations were evident between admission scores and their respective change scores on four measures: the FIM ($r = -0.3419$, $p = 0.015$); LHS ($r = -0.5380$, $p = 0.0001$); SF36 physical component ($r = -0.3920$, $p = 0.006$); and GHQ ($r = 0.6258$, $p = 0.0001$). No significant relationships were observed between the admission scores and change scores of different outcome measures.

iii) Associations between admission scores and the duration of carryover

These relationships were explored by correlating admission scores on each outcome measure with the summary measures. No significant relationship was observed between any of the summary measures and: the EDSS ($r = -0.0421 - 0.0839$, $p > 0.05$); the FIM ($r = -0.1955 + 0.1134$, $p > 0.05$); and the LHS
(r = -0.2522 - +0.1296, p > 0.05). Moderate correlations were observed between:
the mental FS scale and the GHQ summary measure (r = -0.3821, p = 0.009);
the SF36 physical (r = -0.3428, p = 0.020) and mental components (r = -0.3250,
p = 0.028) and their respective summary measures; the GHQ admission score
and the LHS summary measure (r = 0.3151, p = 0.033). A weak relationship was
observed between the GHQ admission score and its summary measure
(r = 0.2908, p = 0.050).

iv) Associations between the magnitude of change made during
rehabilitation and the long term maintenance of change

These relationships were explored by correlating the change scores (between
admission and discharge) and the summary measures. A strong correlation was
observed between the change score and summary measure on the mental health
component of the SF-36 (r = 0.7334, p = 0.0001). There were moderate to
substantial correlations between the following change scores and their respective
summary measures: FIM (r = 0.4934, p = 0.0001), LHS (r = 0.5311, p = 0.0001),
SF36 physical component (r = 0.6276, p = 0.0001) and GHQ (r = 0.5168, p =
0.0001). No statistically significant relationships were observed between the
change scores and summary measures of different outcome measures.
5.5 Discussion

In determining the effectiveness of inpatient rehabilitation there are two equally important aspects to measure: the benefits gained during the inpatient stay; and the carry-over of these benefits to the home environment. The vast majority of the earlier studies have focused on measuring the benefits gained in the short term. Few have investigated whether the gains made during inpatient rehabilitation were sustained over time and in the patient’s own environment. None have documented the pattern or the extent of these changes in any detail. This prospective longitudinal study begins to address these issues by monitoring a range of outcomes at three monthly intervals for one year following the patients discharge into the community. This study is primarily exploratory in nature. Its purpose is to systematically observe the pattern of performance and the length of time changes are maintained in the longer term, rather than to chart precisely the effectiveness of rehabilitation.

5.5.1 Findings

The results of this study showed that positive changes were made during a short period of inpatient rehabilitation, in the outcomes of disability, handicap, HRQOL and emotional well-being. These changes were maintained, in part, in the longer term following discharge, despite worsening neurological status. The time series data showed that the pattern of performance and the duration of carryover differed in each of the outcomes and was variable between individuals.
The progressive nature of MS means that a deterioration in neurological status is the norm (Thompson 1996a; Weinshenker et al 1996a). Without any intervention one could therefore expect to see a reduction in the ability of some patients to undertake everyday activities, over a twelve month period. The results of this study suggest that a short period of inpatient rehabilitation was associated with a delay in the deterioration of these everyday activities. As expected, inpatient rehabilitation did not impact significantly upon impairment. Rather, the improved outcomes were achieved within the limits of the existing impairment. While the mechanism underlying these changes is unknown, it is presumably because of the symptomatic management provided by the rehabilitation programme.

A large amount of information has been gathered from this study. The discussion addresses the findings in relation to each outcome variable in turn. General points are then raised and the study methodology evaluated.

5.5.2 Outcome and summary measures

i) Neurological status

Patients selected for admission to inpatient rehabilitation are often those in the more advanced stages of the disease, with a range of complex interrelated disabilities (Thompson 1996b). They are frequently referred for rehabilitation because of recent deterioration in their condition (Johnson and Thompson 1996). The study population reflected this. It was comprised of patients in the progressive phase, all of whom were moderately to severely disabled, as demonstrated by the high FS and EDSS scores. As expected, the overall
neurological status of this group declined over the twelve month study period. Characteristic of MS were the observed variations in the rate and pattern of the disease course.

A number of epidemiological studies have described the natural history of MS in terms of the time taken to reach a specified score on the DSS or EDSS (e.g. Miller et al 1992; Weinshenker et al 1989, 1996a). Similarly, change in EDSS score has been the principal end point in many large multi-centre drug trials (e.g. IFNB Study Group 1995; Johnson et al 1995; Jacobs et al 1996). Consequently there is a growing body of literature which documents the average rate of change in EDSS scores over the short term. Although this bank of data is not yet sufficient to provide standards for uncontrolled studies of outcome (Quick and Schapiro 1996), it is nevertheless of interest to use the available information to compare the rate of disease progression in this group to others.

Recently, Quick and Schapiro (1996) compiled and analysed the EDSS and DSS data from 37 international studies. Data was included from epidemiological studies and the control and pre-treatment groups of experimental trials. Ten of the studies contained time-series EDSS data for individual patients over a 12 month period. This information has been amalgamated to provide an overview of the number of individuals progressing from one EDSS category to another during a 12 month period (Table 5.13).
Table 5.13 Numbers of individual patients progressing from one EDSS category to another during a 12-month period

| Original EDSS | 0   | 0.5 | 1   | 1.5 | 2   | 2.5 | 3   | 3.5 | 4   | 4.5 | 5   | 5.5 | 6   | 6.5 | 7   | 7.5 | 8   | 8.5 | 9   | 9.5 | 10 | Total |
|---------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 0             |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 3   |
| 0.5           |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 0   |
| 1             |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 0   |
| 1.5           | 1   | 105 | 20  | 4   | 3   | 1   | 1   | 1   |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 136 |
| 2             | 8   | 5   | 2   | 5   | 1   | 2   |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 23  |
| 2.5           | 3   | 1   | 1   | 4   | 2   | 1   | 4   |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 17  |
| 3             | 3   | 1   | 1   | 4   | 8   |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 15  |
| 3.5           |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 21  |
| 4             |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 2   |
| 4.5           |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 26  |
| 5             |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 9   |
| 5.5           |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 108 |
| 6             |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 74  |
| 6.5           |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 52  |
| 7             |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 12  |
| 7.5           |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 18  |
| 8             |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 6   |
| 8.5           |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 4   |
| 9             |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 4   |
| 9.5           |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 2   |
| Total         | 0   | 0   | 1   | 113 | 32  | 8   | 11  | 10  | 23  | 2   | 13  | 8   | 101 | 72  | 58  | 24  | 28  | 15  | 9   | 5   | 0   | 533 |

(Table reproduced with permission from the authors, Quick and Schapiro 1996)

Numbers in bold italics indicate the number of individuals who remained stable throughout the 12 month period
This data has been used to crudely compare the rate of progression of individuals in this longitudinal study to those in Quick and Schapiro’s survey (Table 5.14a,b). To do this patients were first classified according to their baseline EDSS level. This was important since the non-linear nature of the scale means that patients progress through different levels on the EDSS at different rates (Noseworthy 1994a; Weinshenker et al 1996a). The percentage of patients who moved one point or greater on the EDSS was then calculated, and patients were categorised as remaining stable, deteriorating or improving over the 12 month period (Table 5.14a).

In comparing the two data sets, a difference in rate of progression occurred for patients beginning at EDSS levels 6.0 and 6.5. A greater proportion of patients in this study deteriorated over the 12 month period compared to the overall MS population. It is difficult to know why this happened. It might simply be a reflection of the natural variability observed in progression rates (particularly as the numbers in each category were very small). It may, however, partly reflect a selection bias in the patients admitted for rehabilitation. For example, clinical experience suggests that both patients and clinicians appear most aware of the need for rehabilitation when walking becomes increasingly difficult, and when the likelihood of wheelchair dependence is imminent - rather than when there is a deterioration of transfers or self care activities. Consequently many of the ambulatory patients admitted for rehabilitation are often very near to, but not quite at, the point of being wheelchair dependent (EDSS = 7.0). These patients are usually struggling against all odds to “keep on their feet”. It is perhaps then not surprising that these patients “tipped over the brink” from walking to wheelchair dependence over the 12 month period.
Table 5.14a
The progression of EDSS scores over the 12 month study period of this longitudinal study (n=48)

<table>
<thead>
<tr>
<th>Original EDSS score</th>
<th>% stable (+/- &lt; 0.5 point)</th>
<th>% improving (&gt; -1.0 point)</th>
<th>% deteriorating (&gt;+1.0 point)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.0 (n=8)</td>
<td>62.5</td>
<td>12.5</td>
<td>25</td>
</tr>
<tr>
<td>6.5 (n=16)</td>
<td>69</td>
<td>0</td>
<td>31</td>
</tr>
<tr>
<td>7.0 (n=2)</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7.5 (n=5)</td>
<td>60</td>
<td>0</td>
<td>40</td>
</tr>
<tr>
<td>8.0 (n=9)</td>
<td>56</td>
<td>33</td>
<td>11</td>
</tr>
<tr>
<td>8.5 (n=5)</td>
<td>80</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>9.0 (n=3)</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9.5 (n=0)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 5.14b
The progression of EDSS scores over a 12 month period as determined by Quick and Schapiro’s survey (n=276)

<table>
<thead>
<tr>
<th>Original EDSS score</th>
<th>% stable (+/- &lt; 0.5 point)</th>
<th>% improving (&gt; -1.0 point)</th>
<th>% deteriorating (&gt;+1.0 point)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.0 (n=108)</td>
<td>82</td>
<td>3</td>
<td>15</td>
</tr>
<tr>
<td>6.5 (n=74)</td>
<td>90</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>7.0 (n=52)</td>
<td>77</td>
<td>6</td>
<td>17</td>
</tr>
<tr>
<td>7.5 (n=12)</td>
<td>58</td>
<td>0</td>
<td>42</td>
</tr>
<tr>
<td>8.0 (n=18)</td>
<td>61</td>
<td>17</td>
<td>22</td>
</tr>
<tr>
<td>8.5 (n=6)</td>
<td>83</td>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td>9.0 (n=4)</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9.5 (n=2)</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Percentages have been calculated from the raw data provided by Quick and Schapiro 1996 (Table 5.13)
Any comparisons regarding the rate of disease progression must be made with caution - Quick and Schapiro's data is compiled from relatively small numbers of patients in a range of settings, and the sample size of this study is small and biased in its selection. Furthermore, EDSS scores cannot be viewed in isolation. For example the type of MS (Runmarker and Andersen 1993), the disease duration (Weinshenker et al 1996a) and the number of affected neurological systems (Runmarker et al 1994) have all shown to be associated with the rate of disease progression in the short term. Despite these limitations such comparisons enable the rate of disease progression to be viewed tentatively in the context of the wider MS population.

**ii) Disability**

The severity and range of disabilities experienced by this sample are typical of other rehabilitation studies (Francabandera et al 1988; Aisen et al 1996; Kidd and Thompson 1997, in press). As has been the case in every outcome study (see Section 3.1), these results demonstrated an overall improvement in disability immediately following completion of the inpatient programme.

Calculation of group statistics and the group's time series plot of performance levels (Figure 5.4) demonstrated that the median FIM score remained above the baseline for the entire study period. The group's performance level peaked at discharge and then gradually declined over the study period, remaining marginally above the initial baseline score at the twelve month assessment. By definition, however, group scores and their pattern of response over time do not describe most patients' experience. They give no information about
variations in either the degree or the duration of benefit for individuals. They also fail to take into account the fact that patients begin from all levels of disability. A key feature of MS is the individual variability and unpredictability of the disease process. It was considered important to address this in the analysis. This was achieved by graphing each individual’s level of performance throughout the study period and by using summary measures to determine the length of carryover.

As expected in MS, marked variations in both the pattern and extent of carryover in individuals was observed (Figure 5.5). It was interesting that a marked deterioration occurred, between consecutive assessments, in ten patients. This can partly be explained by the fact that it coincided with a relapse in five patients and with an appendectomy in one case. It is worthwhile considering why such a relatively rapid and marked deterioration occurred in the other four cases. A profile of these cases (refer to Appendix 14) shows that although a relapse had not been reported, a deterioration in neurological status had occurred in each of the patients, as indicated by an increase in either the EDSS or FS scores. Perhaps these patients were at a “threshold of disability” where even a small deterioration in neurological status impacted heavily on the overall level of disability. For example, a relatively minor increase in ataxia may result in lower scores in multiple FIM items including upper and lower limb dressing (buttons or buckles), bladder management (the ability to undertake intermittent self catheterisation), and feeding (cutting up tougher foods). Hence as a result of even a relatively minor neurological change, patients may suddenly and markedly reduce their total FIM score. Additional factors also
may have contributed to the reduction in functional ability: two of the four patients were hospitalised, and the other two had moved to different accommodation. These results demonstrate that disability need not progress in a graduated fashion. They support the view that the evolution of disability is complex with many interrelated factors contributing to its progression (Wade 1996).

The summary measures showed that, on average, disability was maintained above the baseline for approximately six months. For the majority of patients who improved their functional performance, the peak was observed at completion of the rehabilitation programme (57% of patients). Following discharge most patients level of performance gradually declined. This was unsurprising for a number of reasons. Firstly, the progressive nature of the disease means that deterioration in function over time is generally the norm. Secondly, the rehabilitation environment is specifically designed to enhance the patients performance: a range of suitable aids and equipment is available; team members are at hand to offer advice and encouragement; and are experienced in facilitating patients (both physically and emotionally) to achieve their optimal levels of performance. Many of these ingredients may not completely transfer to the home environment. Thirdly, patient's often reported that "even though they could perform the task it was completed by the carer(s) because it was quicker to do so". The scoring method of the FIM requires that the patient is assessed on performance not capability. The scores therefore reflected what the patient actually did (or were allowed to do), in the community rather than what they were capable of achieving. This may further
explain why higher scores were achieved in the rehabilitation environment
where circumstances allow the patient sufficient time to undertake tasks more
slowly, but with less assistance.

iii) Handicap

The paucity of research in MS rehabilitation, together with the variation in
measurement instruments used, makes it difficult to comprehensively compare
the handicaps of this study population to others. Some common features
however are apparent. For example, the areas where this population
demonstrated the greatest levels of disadvantage (mobility, occupation and
physical independence), are also the areas where other rehabilitation studies
(Feigenson et al 1981) and population surveys (Cervera-Deval et al 1994;
McLellan et al 1989) report the greatest problems. Similarly the level of
unemployment in this sample was equivalent to those of other moderately to
severely disabled groups (McLellan et al 1989; Harvey 1995).

Calculation of group statistics demonstrated that the mean LHS score remained
above the baseline value for the entire study period. The time series plot of
these scores (Figure 5.6) indicated that the group's performance level peaked
at discharge, and then slowly declined over the study period, remaining only
marginally above the initial baseline score at the final two assessments.

As expected the group response curve did not describe most patients’
experience. The individual time series graphs demonstrated variability in the
pattern and extent of carryover in individuals (Figure 5.7). It is interesting that
the majority of patients (77%) who improved their overall level of handicap peaked once they had returned into the community. This is in contrast to disability where the majority of individuals scored highest at discharge. This may partly reflect the time taken for the social services to undertake the recommendations made by the rehabilitation team. It highlights the relatively long period of time it can take for some rehabilitation outcomes to become apparent and supports the argument for longer term community based studies (Fuhrer 1987; Roy et al 1992; Harwood and Ebrahim 1995).

The summary measures showed that, on average, handicap was maintained above the baseline for approximately six months. The range in the length of carryover was large. Some patients deteriorated from the initial assessment whereas others maintained an improved level throughout the entire study period.

One of the key features of handicap is that it describes the disadvantages experienced by the individual, within their customary physical environment, which limits their "normal" role (Harwood et al 1996). Although two previous studies have systematically investigated the effects of inpatient rehabilitation on handicap (Kidd et al 1995; Kidd and Thompson 1997, in press), only one of these studies measured it after the patient had returned to their "normal" environment. This was at three months following discharge (Kidd and Thompson 1997, in press). Kidd and Thompson's study showed that while handicap was improved in 47% of patients immediately following completion of the rehabilitation programme, a further 30% (total patients improved 77% )
reported improvements at review three months later. The study undertaken in this thesis demonstrated that additional improvements continued to occur, in some individuals, over even a longer period of time. This confirms the view that many handicap issues may take some time to resolve (Fuhrer 1987; Harwood and Ebrahim 1995), and further supports the need for longer term studies.

Another key feature of handicap is that it takes into account the physical environment and the availability of help and resources (Stewart et al 1995). Handicap issues are frequently resolved, at least in part, by the provision and installation of aids and equipment, and the set-up of additional services within the home. As a consequence detailed community care plans and follow-up arrangements are an essential responsibility of the inpatient rehabilitation team (Bakheit 1995). Once recommendation’s have been made, however, their undertaking becomes the responsibility of the community care team. It is therefore of considerable interest to see that the majority of recommendations were commenced within three months of discharge from the unit (refer to Table 5.6). While overall these figures are quite positive, it is notable that one third of the home adaptations took six months or longer, rehousing took twelve months in two of the three cases, and specialist services (such as a cognitive therapy or pain clinic) took twelve months or longer to commence in 20% of cases. It should be noted that these figures indicate the commencement rather than the completion of the home adaptations. Nevertheless this data is perhaps more encouraging than one would have expected given the current level of dissatisfaction with the community care system (Campion 1996; Thompson et al 1997, in press).
It is interesting to know how the level of formal community services utilised by this group of patients compares to the general MS population. Unfortunately few studies have gathered this type of information (Prouse et al 1991; Phillips 1995). Limited comparisons can be made between the results of this study and the community survey undertaken by McLellan and colleagues (1989).

Comparisons are restricted to those with the moderately to severely disabled sub-group (EDSS 5.0-9.5) in McLellan’s survey. Some differences were observed in the use of specific services. For example, 30% of patients in this study were in regular contact with a community nurse in comparison to 53% in the Southampton survey; 40% of this group were receiving community OT in contrast to 16% in Southampton. In other services the utilisation appeared more comparable. For example 28% of this sample were receiving regular physiotherapy compared to 33% in the Southampton group. These comparisons should be interpreted with caution: different methods of data collection were utilised; and the six year time period between the studies means that there may have been some changes in service provision. The discrepancies do, however, draw attention to potential inequalities in the provision of community and social services in different geographical areas of the UK. This is a frequent complaint of patients, carers, and neurological charities (BSRM Working Party Report on MS 1993; MS Society Standards in Health Services for People with MS 1996), and has been a common finding in a number of studies (e.g. Bax et al 1988; Williams and Bowie 1993; Thompson et al 1997, in press).
iv) Health-related Quality of Life

Health related quality of life data is difficult to interpret. Unlike impairment and disability, where a relatively precise quantification of performance can be undertaken, the definition and measurement of HRQOL is far more complex (Whiteneck 1994). This is further complicated by the fact that an individual's quality of life is open to influence from a wide variety of social and environmental issues including family events and financial circumstances (Pfennings et al 1997, in press). Finally, since no previous studies have evaluated the long term impact of HRQOL following rehabilitation, comparisons between other groups cannot be made. The proposed explanations for the observations are therefore purely speculative. Despite these limitations the information gathered is useful in providing pointers and comparisons for future work.

As expected, compared to normative values for the British healthy population this group of moderately to severely disabled patients reported lower levels of health in every single SF-36 dimension (Freeman et al 1996b). Similarly, the physical and mental health summary scores were well below the normative values of the general US population (normative mean = 50 for both scales, Ware 1994a).

The group statistics demonstrated that the physical and mental health components of the SF-36 behaved differently. On average, the mental health summary scores remained above the baseline values for the entire study period. In contrast the mean physical health summary scores fell below the
baseline at the nine month assessment. Differences were also evident in terms of the group's pattern of response over the study period (Figure 5.8 a,b). A substantial improvement was reported in physical health at discharge. This level remained reasonably consistent over the following six months but then declined markedly at nine months. In contrast, the group reported a steady improvement in mental health during the first half of the study period, followed by a decline in the latter six months. These differences are not surprising. The two components are designed to measure different concepts (Ware et al 1995), and it is expected that different outcomes may take different times to become apparent (Fuhrer 1987).

The individual time series graphs demonstrate the variability in both the pattern and extent of carryover in individuals (Figure 5.9 a,b). It is notable that 17 patients (37%) reported lower levels of mental health on discharge from rehabilitation. A review of the questionnaire may partly explain this. The items of the mental component include aspects of social functioning over the previous month which are likely to have been directly limited by the inpatient setting.

A distinctly different pattern is noted for physical health. In this component no-one reported reduced levels of health (compared to the baseline score) until after the six month assessment. This observation does not appear to relate to any of the other outcomes measured. It is somewhat puzzling and difficult to explain. It may be that aspects of the overall "rehabilitation experience", which remain undefined and unmeasured, are accountable. Of further interest is the
relatively sudden decline in the reported levels of physical health following the six month assessment. This may relate to the fact that, up until six months, the levels of disability and handicap had been maintained above the baseline level of performance. Given this decline in disability and handicap, it is perhaps not surprising that a reduction in the reported levels of HRQOL followed soon after.

v) Emotional well-being and Well-being

On admission 74% of patients in this sample demonstrated emotional disturbance (as indicated by a score of five points or more on the GHQ). This level is considerably higher than in other studies where levels vary from between 40-50% (e.g. Rabins et al 1986; Vickrey et al 1995). As previously discussed, this may be partly accounted for by the inflation of scores due to somatic symptoms (Wade 1992; Minden and Schiffer 1990), and the severe levels of disability experienced (Borgel et al 1992; Devins et al 1993). Other factors may include the inherent referral and selection biases of this population. Clinical experience suggests that patients with both emotional and physical problems are more likely to be referred and admitted to inpatient rehabilitation because of the nature and complexity of their problems.

Calculation of group statistics demonstrated that the median GHQ scores remained below the baseline value for the entire study period, indicating sustained improvements in emotional well-being following rehabilitation. The group’s time series plot (Figure 5.10) showed that the lowest level of emotional distress was reported at discharge. The median scores deteriorated slightly at three months, but then remained reasonably stable for the remainder of the
study period. No rehabilitation studies have monitored emotional well-being in the long term and therefore comparisons with other groups cannot be made. However, Jonnson and colleagues (1996) investigated the changes which occurred between admission and discharge from an inpatient rehabilitation programme. The findings of this study at discharge are in agreement with theirs.

The improvements in emotional well-being were sustained throughout the entire study period for the majority of individuals (Figure 5.11). By the 12 month assessment the percentage of patients demonstrating emotional disturbance had reduced from 74% to 48% - an equivalent level to many other studies (e.g. Joffe et al 1987; Rabins et al 1986; Vickrey et al 1995). It is interesting to postulate why this occurred. Previous studies have demonstrated that emotional distress is associated with physical disability (Borgel et al 1992; Devins et al 1993). Perhaps the reduction in physical disability could partly account for the improvements observed in emotional well-being, at least in some individuals. Alternative explanations may also exist. For example, a number of studies have demonstrated that emotional distress is associated with perceived control over health outcomes (Wassem 1991; Shnek et al 1995). It is possible that the educational and supportive processes of rehabilitation empowered the person to become more actively involved in their health care and welfare needs (La Rocca et al 1993; MacLaren 1996). It might also be that the additional and regular attention received through participation in this study may have resulted in an improved level of emotional well-being. Either of these latter two suggestions could help to explain why the improvements were
maintained throughout the entire study period despite a gradual deterioration in the disease process.

Fluctuations in emotional well-being were seen in individuals throughout the study period (Figure 5.11). It is of interest to know whether these variations are within the normal range for individuals over time. Unfortunately no studies provide this type of information. A study by Dalos and colleagues (1983) used the GHQ-28 to monitor emotional well-being at monthly intervals over a twelve month period. They observed that an increase in emotional disturbance was nearly always associated with increasing disease activity. The findings of this study differ from theirs. Of the 15 patients who reported a total of 22 relapses, an increase in GHQ scores was observed on only nine occasions. These differing results may be partly explained by variations in the two samples. This group was comprised solely of people in the progressive phase, whereas Dalos's patients were in either the relapsing/remitting or the progressive phase of MS.

5.5.3 Correlation analyses

The relationships between pathology, impairment, disability and handicap are poorly understood (Wade 1996; Newham 1997). Earlier work in rehabilitation has focused on the measurement of one level of outcome - disability. It has been suggested, however, that disability may be of less concern to patients than other areas such as handicap and quality of life (Ware 1993a; Harwood and Ebrahim 1995). It is important to study the relationships between different outcomes to determine whether measurement of a single entity, such as
disability, can be a valid representation of the effectiveness of rehabilitation, or whether a wider range of outcomes is needed. Furthermore, observing what happens to different dimensions over time may increase our understanding of how disability and handicap progress and thereby help in the formulation of interventions according to need (Whiteneck 1994).

The correlation analyses demonstrate that the associations between the different outcome variables range from weak to moderate indicating that each of these scales measures a different concept (Table 5.12). In other words, changes in one level of measurement cannot be extrapolated to others. This endorses the argument that outcome measurement at multiple levels is necessary to ensure a comprehensive description of the impact of disease and treatment (Hobart et al 1996a; Harwood et al 1996; Newham 1997).

This correlation data supports clinical experience. For example, that functional improvements can occur despite irreversible impairment (Pope 1992; Edwards 1995; Wade 1996), and that handicaps can be reduced even when impairments and disabilities remain unchanged (Stewart et al 1995; Harwood et al 1996). It confirms the clinical impression that emotional status can be related to aspects such as handicap and HRQOL (Wade 1992).

The individual nature and wide ranging symptoms of MS ensure that clinical judgement will always play a major role in the selection of patients for rehabilitation. However, additional objective information concerning which patients are most likely to benefit from intervention is likely to be invaluable. To
date, only a few studies have investigated which factors influence rehabilitation outcome (e.g. Carey et al 1988; Langdon and Thompson 1995; Grasso et al 1996). Carey and colleagues (1988) demonstrated an association between the level of functional status on admission and the functional benefits gained during rehabilitation. Langdon and Thompson (1995) found that three variables were related to a reduction in disability following rehabilitation: the admission level of disability, vocabulary skills and cerebellar function. Grasso and colleagues (1996) reported that the EDSS admission level related to the benefits gained in disability. The results of this longitudinal study suggest that, in the main, benefits gained are only significantly associated with the admission score on that specific variable. In terms of maintenance of benefits, the patients who made the greatest change during the inpatient stay were also those who maintained improvements the longest, but only with respect to the same variable. There are no clear relationships between different outcomes and carryover of benefit.

5.5.4 General comments

Some general comments can be made with regard to the study's findings. Firstly that the overall changes measured for the group were relatively small in each of the outcomes. The distribution of these changes however was wide, suggesting that patients who improved were balanced by others who deteriorated. This "evening out" which occurs in group statistics may give the impression that rehabilitation is associated with only small changes. The individual time series graphs show that this is not necessarily the case. Marked clinical improvements can occur, at least in some individuals. This information
can aid the identification of who does and who does not benefit from rehabilitation, thereby enabling better targeting of resources in the future.

Secondly, by reviewing patients at regular intervals it was questioned whether a pattern would emerge in the carryover of benefits. It was anticipated that this information would prove useful in developing guidelines for patient review. Currently no such guidelines exist and hence decisions are made on an ad hoc basis. The summary measures showed that, on average, the benefits gained during rehabilitation were maintained for approximately six months. This suggests that it may be relevant to review patients six months following discharge. The individual graphs, however, clearly highlight that from an individuals perspective, rigid guidelines for review are inappropriate. Hard and fast rules can not account for the marked variability seen in MS patients. This supports the argument by both professionals and patients that MS services must be flexible and responsive to the needs of the individual (Thompson 1996c; Thompson et al 1997, in press; MS Society Standards in Health Services for people with MS 1996).

Thirdly, the correlation analyses demonstrated that the size of the benefits gained during rehabilitation were significantly related to the length of time these changes were maintained in the longer term. This suggests that if you can predict in advance who will benefit from rehabilitation, you may be able to predict carryover.
5.5.5 Evaluation of methodology

i) Study design

As previously discussed this study was exploratory in nature. Its primary intention was to systematically observe both the pattern of performance and the maintenance of benefits, in a variety of outcomes, in the year following discharge from inpatient rehabilitation into the community. For this purpose a single group, pre- and post design with assessment at multiple time points was used. Ideally a control group would have been incorporated so that causal inferences about the effectiveness of rehabilitation could have been made (Pollock et al 1993; Moses 1995). Ethical and logistical considerations, and the expectations of patients and families, precluded a randomised controlled study in this situation. As a consequence causal inferences cannot be made since several different relationships could have produced the change in outcomes (D'Agostino and Kwan 1995). This study was therefore undertaken in parallel with a randomised controlled trial, the methodology and results of which have been discussed in the previous chapter.

No trial, particularly if it is longitudinal in design, can be isolated from extraneous factors (Pollock et al 1993). Over time the patients are exposed to more and more undefined causal factors, making rehabilitation an increasingly small part of the totality that influences them (Johnston 1987). These issues are further compounded by: (1) the variable, unpredictable and generally progressive nature of the disease course; (2) the fact that many of the outcomes measured are influenced by social and environmental factors other than treatment (e.g. social and community support). These confounding factors
make interpretation of the data difficult. Some of the key confounding variables were therefore recorded and monitored. These included: disease severity; number of relapses; rate of re-hospitalisation; changes in drug intervention; and in formal community support. Any hypotheses about the impact that these factors may have had on the long term carryover of changes can only be made in very broad and speculative terms.

Choices about the number of observations and their temporal spacing can have important effects on the conclusions of a study (Collins and Johnston 1995). While guidelines exist for the frequency and timing of follow-up assessments in stroke (Stroke Task force 1990) no such recommendations are available for MS rehabilitation research. Based on clinical experience, and in line with the number and spacing of assessments suggested by the Stroke Task Force, observations were limited to six sequential time points. Significant fluctuations may have occurred between these assessments. These are unknown.

Despite the considerable effort made in gathering a complete set of data, information was missing for six patients at one or more time points. Missing data is a problem common to all longitudinal studies (Holford 1992). The average follow-up rate has been estimated at between 60-80% (Forer 1987). The follow-up rate of this study (88%) was therefore considered to be high. There are no strict rules as to how to deal with missing information. It is a matter of judgement whether to estimate the data, or whether to ignore the patient records with missing values (Uberla 1981). In this study details of all patients whose data was incomplete were recorded. A review of this
information showed a systematic bias - virtually all of these cases were severely disabled. Consequently, rather than risk biased inferences by ignoring this data all available information was included at each time point.

It is appreciated that the generalisability of this study is limited: (1) all subjects were in the progressive phase; (2) all were moderately to severely disabled; (3) the data was acquired solely from one service rather than from multiple sites.

There is no assurance that the observations from this study will apply to groups free of these selection biases (Nelson et al 1988). It is noted, however, that the baseline scores for each of the outcome variables were very similar to the randomised controlled trial. This can only be partly explained by the overlap of patients (n=26) in both studies. It also suggests that the characteristics of the patients selected for rehabilitation at the NRU are reasonably constant. This is further supported by the results of regular audits (Freeman et al 1996a,d) and previous studies (Kidd et al 1995; Kidd and Thompson 1997, in press). The characteristics of this population are also similar to those of other rehabilitation studies (e.g. Francabandera et al 1988; Brosseau and Wolfson 1994; Grasso et al 1996; Aisen et al 1996). This goes some way to support the generalisability of these results to this service and others.

iii) Interpretation of the outcome measurement data.

The analysis of outcomes in this study was restricted to the total scores on each measure. The total scores are derived by summing a number of different items on each scale (13 items on the FIM, six dimensions on the LHS, 28 items on the GHQ, and 36 items on the SF-36). An identical score can therefore
result from improvement or deterioration in a range of different items. As a consequence, interpretation of the serial data is limited to describing the outcomes in overall terms and not the individual profile. It is not known which specific aspects of outcome altered during the study period.

The measurement instruments chosen were identical to those used in the randomised controlled trial. The problems identified with their use and interpretation also relate to this study (refer to section 4.6.2). An added difficulty in longitudinal trials is the issue of repeated measurements (Pfenning et al 1995). The effects of repeated measurements on the instruments used in this study, however, have not been previously documented.

Finally, but by no means of least importance, it is crucial to determine what is considered a successful outcome of MS rehabilitation. There is a general agreement that the main emphasis of interventions for chronic and progressive diseases is to maximise function, to prevent secondary complications and to maintain the best possible quality of life within the limits of the disease (Edwards 1996). Positive gains, such as reduced disability, are relatively easy to measure in the short term. It is far more difficult to demonstrate aspects of long term success, such as the prevention of secondary complications (e.g. pressure sores or contractures) or the delay in deterioration of outcomes. This is especially problematic when the disease is progressive. In defining the success of rehabilitation a number of questions need to be answered. For example, is rehabilitation only considered to be effective if aspects of daily life are maintained for a specific (currently unknown) length of time?,
additional, albeit small, improvements expected to continue following discharge into the community? Should there be different criteria for success at different stages of the disease? Indeed does the unpredictability and variability of the disease course prevent valid measurement in the longer term? Consideration of these complex issues is essential if data from longitudinal studies is to be meaningfully interpreted.

5.6 Conclusion

Little is known about the long term benefits of inpatient rehabilitation. This longitudinal study has systematically measured the levels and the carryover of benefits, in a range of dimensions, for one year following discharge from rehabilitation. The results showed that positive changes were made during a short period of rehabilitation in the outcomes of disability, handicap, HRQOL and emotional well-being. These changes were maintained, in part, for approximately six months following discharge, despite worsening neurological status. The duration of carryover differed in each of the outcomes and was variable between individuals. It was observed to decline over time. These results reinforce the widely held belief, of both professionals and patients, that a continuum of care from the inpatient hospital setting to the community is essential in the overall management of patients with MS (e.g. BSRM Working Party Report on MS 1993; Campion 1996; MS Society Standards in Health Services for people with MS 1996).

Like most longitudinal studies, this study has posed more questions than it has provided answers. It does however serve as a useful starting point for
increasing our understanding of the dynamics of disability, handicap, HRQOL and emotional well-being, and the relationships these have to the outcome of rehabilitation. Just a small proportion of the wealth of information generated from the study has been evaluated in this thesis. This information begins to provide important pointers and comparisons for future rehabilitation research.
Chapter 6
Conclusions

Great strides have been made in our understanding of multiple sclerosis in recent years (McDonald 1995). This has led to the recent development of a number of new disease modifying pharmacological therapies which have been shown to reduce relapse rate (IFNB Study Group 1995; Johnson et al 1995; Jacobs et al 1996). However, while these advances offer renewed hope to MS patients a clinically meaningful effect of drug therapies on disability has yet to be demonstrated (Thompson and Noseworthy 1996); the aetiology of MS remains unknown; and neither a prevention nor cure is likely in the near future (Compston 1994; Johnson 1996). Progressive disability remains the experience of most individuals, and symptomatic management and the provision of supportive services continues to be the cornerstone of health care intervention. As a consequence there is an increasing recognition of the need to address disability, handicap and quality of life in both the provision and evaluation of health and social service interventions (Runmarker and Andersen 1993; Rodriguez et al 1994; Devinsky 1995; Johnson 1996; Hobart et al 1996a).

Inpatient rehabilitation is advocated as an important intervention in the overall health care management of MS (ABN Working Party 1992). Frequent demands to increase service provision in this area are made by patients, health care professionals and the neurological charities (Neurological Provision-key areas and targets 1991; BSRM Working Party Report on MS 1993; MS Society Standards in Health Services for People with MS 1996). Inpatient rehabilitation
is however costly, its effectiveness has not been established, and little is known about the long term carryover of benefit. This information is essential in enabling decision making that is evidence-based and ensuring continued improvements in patient management (Rosenberg and Donald 1995; Sackett et al 1996). Furthermore, it has been argued that a lack of evidence may potentially threaten the continuation of a service in this era of evidence-based medicine, where the emphasis is on targeting resources to interventions of proven benefit (Grimley-Evans 1995; New 1996). In response to this need, this thesis has evaluated the effectiveness of inpatient rehabilitation for patients in the progressive phase of MS. It has focused on the measurement of outcome as a, now widely accepted, means of determining effectiveness (Frater and Costain 1992; Wade 1993; Sackett et al 1996).

6.1 Tackling the methodological dilemmas encountered

Developing a methodologically sound study design, which remains clinically relevant and practical, is essential for results to be meaningful, but has proven extremely difficult to achieve. As detailed in the introduction (Sections 3.2 and 3.3), a number of methodological problems were encountered, and some remained unresolved, in the design of previous studies. The resultant flaws have markedly limited the interpretation of this earlier work, and have failed to provide conclusive evidence of effectiveness. The basis of this thesis is the design of two closely linking outcome studies which have achieved methodological improvements in both the content and format of earlier studies.
These improvements include:

1. the use of a universally accepted theoretical framework upon which to conceptualise the goals and outcome measures of intervention - the WHO's International Classification of Impairments, Disabilities and Handicaps

2. the use of prospective study design in each of the studies

3. the incorporation of a control group by the use of a stratified randomised wait-list controlled design, to enable causal effects to be determined

4. the incorporation of two different, but complementary, study designs, to investigate, in parallel, the effectiveness of rehabilitation in both the short and long term

5. selection criteria to ensure a relatively homogenous study population, and to reduce the likelihood of change due to spontaneous neurological recovery

6. the regular and systematic monitoring of neurological status throughout the study duration to facilitate meaningful interpretation of the results

7. the use of standardised outcomes measures which were appropriate to the goals of the rehabilitation programme, and whose scientific and clinical properties have been evaluated
8. a broadening of the scope of assessments to include both physician-oriented and patient-oriented measures, thereby providing a more comprehensive and relevant evaluation of the wide-ranging goals of multidisciplinary inpatient rehabilitation

9. the use of single assessors throughout the study period to increase the reliability of results

10. the continued assessment of outcome following discharge from rehabilitation into the community, thus enabling the longer term maintenance of change to be explored

These methodological improvements mean that, compared to earlier work, the results of the studies undertaken in this thesis can be interpreted with more confidence.

6.2 The methodological challenges that remain

A number of methodological problems remain unresolved in the design of clinical trials of MS inpatient rehabilitation. These include:

1 Lack of blinding

Lack of blinding of either the patient, the treating clinician(s) or the investigator risks false positive findings concerning effectiveness (Noseworthy et al 1994b). Practical and financial considerations prevented this from being incorporated
into the design of these studies - although not involved in treating any of the patients, both the assessors were based at the rehabilitation unit.

2 Lack of a placebo-control group

The unpredictability of MS and the tendency to spontaneously recover following relapse suggest that a placebo group should be included in studies of effectiveness (Whitaker et al 1995). The lack of a placebo-control group does not enable a change due to aspects such as increased attention or the environment to be distinguished from a change due to the intervention. It increases the risk of false positive findings. This ideal is, however, difficult to achieve for a number of reasons. Firstly, the active ingredients of the rehabilitation process have not been identified either individually or in combination. The placebo condition is therefore, as yet, undefined. Secondly, patients in the placebo group would have to be admitted to a "rehabilitation style setting" for approximately three weeks where they would participate in activities designed to have no effect on disability, handicap, HRQOL, or emotional well-being. In addition, both the patients and the staff would need to be unaware that this was the sham intervention. From both a practical and economic perspective, it is difficult to see how this could be incorporated.

3 Limited generalisability

As discussed in previous chapters, the results of single centre studies are limited in their extrapolation to wider settings where the techniques, patients and environmental factors may be different. For definitive conclusions to be
made these results need to be confirmed across settings, investigators, and study populations.

4 Lack of a control group for the longitudinal study.

Ethical considerations and the expectations of both patients and clinical staff strongly influenced the decision not to utilise a randomised controlled design in the longitudinal study. It is, however, acknowledged that these factors may be considered insufficient justification for precluding a controlled study design (Newham 1997). The lack of a control group means that causal inferences about the ability of inpatient rehabilitation to effect a long term change cannot be made (Senn 1994; Moses 1995). Interpretation of the longitudinal data is limited to theoretical conjecture based on the study observations.

5. Lack of a detailed description of the rehabilitation process

The intention of these studies was to focus on outcome not process and hence the lack of a detailed description of the specific rehabilitation interventions is not a fundamental methodological problem. It does, however, make it difficult to replicate these studies in the future. This has been a common problem to many rehabilitation studies (Wade et al 1992) for a number of reasons. Firstly the recognition that a key feature of the rehabilitation process is the individualised treatment approach to each patient's unique set of problems (Freeman et al 1996c). Secondly, that the treatment and goal setting process is inherently dynamic in nature, responding to the patient's current problems, and adapting treatment to changes as they occur (Partridge and Edwards 1996). Thirdly that the rehabilitation package is comprised of input from a number of different
professionals, each at varying levels of clinical experience, and perhaps with a different emphasis or approach to treatment. Although these combined factors make it an extremely difficult and onerous task, it should be possible to describe the interventions in sufficient detail to enable replication of the study.


The unpredictability of the disease course and the diversity of clinical manifestations in MS mean that definitive studies are likely to require large numbers of patients (Goodkin 1996). Small numbers increase the chance of false negative conclusions (type two errors). They also make it difficult to allow valid subgroup analysis (e.g. by severity of specific impairments) so that types of patients likely to benefit from inpatient rehabilitation may be identified.

6.3 The findings of these studies

The main objective of MS rehabilitation is to lessen the impact of the disease on daily life, and enable the patient to realise his or her own potential within the limitations posed by their disease (Mertin 1994; Thompson et al 1994). The results of both the randomised controlled study and the longitudinal study demonstrate that this objective can be successfully achieved. The results of the randomised controlled trial show that a short period of multidisciplinary inpatient rehabilitation was effective at reducing disability and handicap in patients with progressive MS, despite unchanging levels of neurological impairment.
The longitudinal study suggests that, on average, these benefits were maintained for approximately six months following discharge into the community. Rehabilitation was able to preserve function, to some extent, even in the presence of disease progression. Furthermore, rehabilitation was also associated with improvements in health-related quality of life and emotional well-being. These studies conclude that patients in the progressive phase of MS, who are identified by an experienced multidisciplinary team as having the potential to benefit from inpatient rehabilitation, are suitable candidates for inpatient rehabilitation and can gain considerable benefit from this intervention.

In addition to corroborating the non-experimental findings of earlier studies, this thesis demonstrates that it is possible to undertake research into MS rehabilitation using a randomised controlled design with objective outcome measures. The longitudinal study demonstrates that carryover of changes in the longer term can be measured. The information provided by each of these studies is complimentary. The two approaches combined help to provide a totality of evidence upon which rational clinical judgements can be based (Black 1996).

6.4 Clinical Implications

A number of service implications have been highlighted by the results of these studies:

1. The need for a continuum of care in the community

Inpatient rehabilitation should not be reviewed in isolation. It is only one part of the comprehensive model of care for patients with MS. The results of these
studies highlight the necessity for a strong link between the rehabilitation setting and community and domiciliary based services. They indicate that continuity of care is required between these settings to successfully achieve and maintain many of the goals of rehabilitation, and to meet the changing needs of the patient in a comprehensive and timely manner. This is in agreement with a number of reports (National Audit Office, Physical Disability and Beyond 1986; BSRM Working Party Report on MS 1993; Williams and Bowie 1993; MS Society Standards in Health Services for People with MS 1996; Thompson et al 1997 in press), and surveys of people with MS (McLellan et al 1989; Perry 1994). The results suggest that there is a need for formal and structured community links to enhance the carryover and maintenance of changes in the longer term. This is crucial in ensuring that inpatient rehabilitation does not add to the current fragmentation of services but plays its part in drawing the elements together into an integrated model of care designed to meet the ever-changing needs of the person with MS.

2 The importance of the selection process in inpatient rehabilitation

It has been argued that the selection process is critical to the efficacy of inpatient rehabilitation (Freeman and Thompson 1997, in press), and that multidisciplinary assessment, to determine who will benefit from this service, underlies its success (Bakheit et al 1996; Thompson 1996c). Currently notable differences exist in the selection process, both within and between different countries. Important influences on this process appears to be the structure and funding mechanism of individual health care systems, and the wealth of the country (RIMS Conference Proceedings 1996). For example, in some countries
patients with MS are routinely offered inpatient treatment on an annual basis (Vaney et al 1996a), whereas in others patients are selectively admitted according to current need (Freeman et al 1996a). Even within the same country significant differences are evident. In the United Kingdom, for example, some services rely on the opinion of the consultant to decide who should be admitted (Fuller et al 1996), while others depend upon a single assessment by an experienced multidisciplinary team (Johnson and Thompson 1996; Bakheit 1996). Although investigation of the selection process has been made in other neurological disorders (e.g. stroke rehabilitation, Poduri et al 1996), no such research has been undertaken in MS rehabilitation. As a consequence the selection of MS patients for admission to inpatient rehabilitation is based solely on clinical judgement.

This thesis did not look specifically at the selection process. Furthermore, the results did not pinpoint specific features which could predict a successful outcome. Admission guidelines are therefore unable to be determined. The studies did, however, record the wide range of benefits gained thereby indicating the types of problems which may be improved by inpatient rehabilitation. Importantly these results illustrate that even very severely disabled patients may benefit in a variety of dimensions. This information is likely to prove especially useful to clinicians who may have considered that patients with severe disability cannot benefit from this intervention (Greenspun et al 1987; La Rocca and Kalb 1992). Secondly, the results will raise awareness that dimensions such as disability and handicap can be improved despite unchanging impairment. The common misconception that changes in
function are dependent upon improvements in neurological status has often meant that health professionals including neurologists, physicians and general practitioners have failed to refer patients for inpatient rehabilitation (Greenspun et al 1987; La Rocca and Kalb 1992; ABN Working Party Report 1992).

The huge variation within and between patients, together with the multiplicity of symptoms and the way in which they interact, means that clinical judgement will always play a major role in the selection process. In our opinion this selection process is most effective when it is undertaken by a multidisciplinary team experienced in assessing and treating patients with MS. Integral to this process is the identification of areas of improvement and the establishment of achievable goals, agreed by both the patient and the team, prior to admission (Freeman and Thompson 1997, in press).

3 Information for purchasers and providers

The NHS’s internal market philosophy, together with the rising costs of health care in recent years, demands the allocation of resources to areas which are of proven value (New 1996). Indeed some now advocate that activities should be continued “only if they are based on proof of benefit through controlled observations” (Johnson 1996). It is therefore important that rehabilitation provides rigorous scientific evidence of its effectiveness if rationing of this service is to be limited. This would seem to be particularly important in this area where (1) many negative preconceptions exist about the nature of MS and the effectiveness and validity of inpatient rehabilitation (Greenspun et al 1987; BSRM Working party Report on MS 1993; Wade 1995; Bakheit 1995);
(2) there is a lack of good scientific evidence on which to base clear purchasing
decisions (La Rocca and Kalb 1992). The randomised controlled trial is the first
controlled study to investigate the effectiveness of MS rehabilitation. It provides
evidence for purchasers that rehabilitation is effective in reducing disability and
handicap in patients in the progressive phase of the disease.

While these studies provide valuable information about outcome, they do not
provide information about the cost, or the cost-benefits of this intervention. This
information is also essential for the appropriate allocation of resources
(Holloway 1996).

4 The need to review patients following discharge

No guidelines exist regarding the optimal time to review patients following
discharge from rehabilitation. Currently this decision is made on an ad hoc
basis. The results of the longitudinal study suggest that (1) review of patients is
necessary since the benefits gained from rehabilitation decline over time; (2)
re-admission for further rehabilitation may be necessary as new problems
arise. The summary measures indicate that, in general, six months is likely to
be the most relevant time for review since it is at this point that performance
has declined to the initial baseline level. It is important however that rigid
guidelines for review are not enforced. The natural variability and
unpredictability of the disease mean that flexibility of review, as well as of
service provision, is a high priority for both patients and professionals (MS
Society Standards in Health Services for People with MS 1996; Thompson
1996b).
6.5 The Future

Despite the complexity of undertaking clinical trials in MS in looking towards the future there are a number of exciting challenges ahead:

1. Improved study designs

More well designed research studies are required to establish the effectiveness of inpatient rehabilitation in MS. There are a number of realistic goals which will further improve the methodological advancements made in this thesis:

(a) Blinding of the assessor

To date, no study has incorporated blinding into its design. Given sufficient resources, single blinding of the assessor should prove relatively easy to incorporate into future clinical trials. This will reduce the likelihood of type 1 errors (Noseworthy 1994b). While double blinding is preferable in theory it is difficult to envision how it could be achieved in practice. Even in drug trials, where this might appear relatively straightforward, the presence of side effects makes “true blinding” difficult (Gore 1981; Daly et al 1991; Noseworthy 1994b).

(b) A multicentre randomised controlled trial

This study has demonstrated that it is possible to utilise a randomised controlled design to evaluate the effectiveness of MS rehabilitation, at least in the short term. The next step is to undertake a multi-centre study to improve the external validity of the results, and increase the sample size. Given the relatively small numbers of specialist MS units, this may necessitate multi-centre participation involving a number of countries. An attempt to achieve this is currently being planned by a co-ordinated network of task groups throughout
Europe (Rehabilitation in Multiple Sclerosis {RIMS} and MS and Rehabilitation Care and Health Services in Europe {MARCH}, Ketelaer and Pacolet 1996). It is important in these multicentre studies that there is standardisation of selection criteria, rehabilitation input and measurement.

(c) A longitudinal study design which involves a comparison group.

Ethical considerations and the expectation of patients, their families and health care professionals mean that it is unlikely that randomised placebo-controlled longitudinal trials will be undertaken. It is, however, both ethical and practical for studies to involve a comparison group participating in a different therapy (e.g. outpatient therapy). This will help to determine whether a particular rehabilitation approach is more effective in the long term maintenance of benefits.

As discussed earlier, methodologically sound study designs, which remain clinically relevant and practical, are essential for results to be useful. In the basic sciences this has led to the creation of a number of International Task Forces specifically formulated to provide recommendations for MS clinical trial designs (e.g. The IFNB Study Group, Paty et al 1995; The Clinical Outcomes Task Force, Whitaker et al 1995, Goodkin 1996, The Clinical Outcomes Assessment Task Force, Rudkin et al 1996). Although multi-disciplinary as well as multi-centre collaboration does occur within the field of MS rehabilitation (e.g. the RIMS and MARCH Task Groups, Ketelaer and Battaglia 1991; the MS Consortium in North America, La Rocca et al 1994), this is relatively limited and has been slow to develop. It is suggested that improved collaboration between
different disciplines (Freeman et al 1996c) and between rehabilitation centres both nationally and internationally, may help to solve some of the methodological and economic difficulties which currently hinder research in this area.

2 The evaluation of cost and cost-effectiveness

While earlier studies placed very little emphasis on evaluating the costs and cost-effectiveness of rehabilitation in MS, there is now an increasing recognition of the need to include this in future studies (Johnson 1996, Holloway 1996; Thompson 1996c). As discussed earlier (Section 1.2), the cost of MS to patients and society is considerable, with estimates ranging from $12769 to $22875 per annum per patient in the USA (Harvey 1995) and up to $1.2 billion per year for the 87000 patients in the UK (Holmes et al 1995). Furthermore, studies have shown a significant association between total health care costs and neurological dysfunction (Harvey 1995; Holmes et al 1995) and disability in MS (Bourdette et al 1993). This suggests that improving self-care and avoiding preventable hospitalisations (e.g. preventing bladder and bowel complications, and pressure sores) may lower the overall health care costs of MS. Since these are two of the key aims of MS rehabilitation it is of considerable interest to determine its cost effectiveness. For the cost analysis to be valid and reliable, multidisciplinary collaboration between a range of professionals including health economists, clinicians, public policy analysts, quality of life researchers and ethicists is crucial (Holloway 1996).
3 Identifying which is the optimal approach

Although there are many common features of inpatient rehabilitation practice, a number of differences exist in both service provision and delivery. One example, referred to earlier, is the variation in the selection process between different rehabilitation units. Others include the discrepancy in the availability of specialist inpatient and outpatient MS facilities in different countries (Paty 1994); and the varying lengths of inpatient stay (Feigenson et al 1981; Kidd et al 1995). Studies are urgently needed to compare the effectiveness of different inpatient rehabilitation approaches, to determine how these services should be best organised (e.g. specialist neurorehabilitation units versus general rehabilitation units), and to compare multidisciplinary inpatient rehabilitation to other models of service delivery such as unidisciplinary, outpatient and community care. While it is acknowledged that studies of this nature are difficult to design, they have proven to be possible in the rehabilitation of other neurological conditions (e.g. Gladman et al 1993; Langhorne et al 1993; Greenwood et al 1994; Drummond and Walker 1995; Forster and Young 1996).

As mentioned previously, an attempt to address these issues is currently being planned by a co-ordinated network of task groups throughout Europe (the RIMS and MARCH task groups, Ketelaer and Pacolet 1996). This group is planning to undertake a multi-centre study to compare the efficacy of three different rehabilitation regimes: (1) intensive multidisciplinary inpatient rehabilitation; (2) maintenance outpatient rehabilitation for two days each week; (3) unidisciplinary treatment in the community. Assessments of impairment, disability and handicap will be undertaken at six monthly intervals over a three
year period to determine the outcome of each model of care. It is anticipated that studies such as this will increase our knowledge concerning the most effective model of care for MS.

4 Focusing on process as well as outcome

The focus of this thesis, and of earlier studies, has been on measuring the outcomes of care to determine the effectiveness of rehabilitation. Some argue, however, that measuring well supported processes of care is more enlightening (Davies and Crombie 1996). They state that the evaluation of process (what is done to patients, where, when, and how) helps to identify specific shortcomings in the service, and thereby point the way towards what must be changed. It is suggested that each approach has its strengths and limitations, and that both provide valuable and complementary information in the evaluation of health care interventions.

A recent method for monitoring and evaluating the delivery of MS rehabilitation care is the Integrated Care Pathway (ICP). The ICP documents the expected interventions occurring during a given episode of patient care. (Rossiter and Thompson 1995a,b). It details aspects of service delivery and patient management (such as number and timing of assessments and treatments, goal achievement), and records when and why variations have occurred from the expected procedure of the rehabilitation process. The use of this instrument, in parallel with the measurement of outcomes, should prove valuable in relating aspects of process to outcome.
As with any measurement instrument, the ICP has its limitations. Importantly, it does not describe or systematically measure the specific types and methods of interventions employed in the rehabilitation process. Evaluation of these aspects of process are of equal importance if the mechanism of the observed changes is to be determined. This will require the description and measurement of specific approaches or techniques utilised (Plant 1996), information which to date has been rarely documented (Partridge and Edwards 1996). Consequently, a major task in the future evaluation of process, will be to reach a consensus on a common clinical language to describe and measure the contents of the rehabilitation “black box” (Wade et al 1992; Gladman et al 1996).

5 Assessing the parts as well as the whole

Comprehensive rehabilitation is comprised of a wide variety of techniques, by a range of disciplines, to address a number of different problems. In order to know what specific components of the rehabilitation package are most effective these entities must be evaluated (La Rocca et al 1994; Plant 1996). Although some trials have been undertaken examining specific interventions (e.g. Vowels and Pelosi 1983; Petajan et al 1996; Fuller et al 1996; Jones et al 1996), these are relatively few in number. Many more are required to inform and improve clinical practice.

6 Evaluating the comprehensive model of care

While there is a clear need for specific interventions of the inpatient package to be assessed, it is also recognised that inpatient rehabilitation units deal with
only a small part of the long term process of rehabilitation. The inpatient stay should therefore not be considered as a distinctly separate phase of care, but as an integral part of a comprehensive model of care that includes community and domiciliary based services (Thompson 1996b). In line with this view, some suggest that studies should be extended to investigate the comprehensive model of care (e.g. the assessment of inpatient, outpatient and community therapy combined as a 'package' of care). Differing opinions exist as to the usefulness of this approach. Those in favour argue that the combined package of care is reflective of current clinical practice and therefore should be evaluated as a whole (Ketalear and Pacolet 1996). Opponents suggest that studies of this nature are fraught with methodological difficulties, and that meaningful interpretation of their results may prove impossible (Plant 1996). Furthermore, the costs and practical difficulties of undertaking such studies are considerable, and perhaps prohibitive in an area where no vested commercial interests lie (Freeman and Thompson 1997, in press).

7 Natural history studies of MS

Because of the changeable and progressive nature of MS over time, information from longitudinal, population-based, studies is necessary to determine whether the fluctuations observed are within the normal range, and whether the rate of deterioration differs from that expected in the absence of treatment. Furthermore, in order to improve patient management in the long term a picture of how the disease progresses, what factors are important in influencing the disease, the effect of any new impairment on function, and the
interactions and development of disability, handicap and quality of life is required (Perry 1994; Wade 1996).

To date, population-based studies have viewed outcome rather narrowly, using measures which have focused on disease severity, relapse rate, lesion load and disability as measured by the EDSS (e.g. Weinshenker and Ebers 1987; Weinshenker et al 1989; Weinshenker 1994; Runmarker et al 1994). The few studies that have investigated broader issues have been limited by their cross-sectional design (e.g. McLellan et al 1989; Rodriguez et al 1994; Midgard et al 1996). As a consequence relatively little information is available about the natural history of MS both in terms of patient-oriented outcomes such as disability, handicap, emotional well-being and HRQOL, and with regard to the risks of secondary complications. This type of information would prove invaluable in evaluating the long term effectiveness of interventions. It is acknowledged, however, that the interpretation of this information may prove more difficult in the future as a greater variety of disease-modifying pharmacological agents are used in ever increasing numbers of patients.

8 The refinement and development of suitable outcome measures

It has been suggested that until suitable instruments are available for use in evaluation, rehabilitation services will not develop by virtue of a lack of information rather than a lack of benefit to the patient (Harwood et al 1996). This is a particular problem in the area of chronic and progressive neurological disease where few measurement instruments are available (Hobart et al 1996a) and even fewer have been comprehensively evaluated (Hobart et al 1996b).
Using poorly evaluated instruments may lead to misleading results, and thereby affect important clinical decisions.

In general, the acceptability, reliability, and validity of instruments has been reasonably well established. However, responsiveness which is thought to be the most important criterion in evaluating effectiveness, has been less extensively examined (Deyo et al 1991; Fitzpatrick et al 1992a; Guyatt et al 1992). Even those instruments developed in accordance with strict psychometric criteria and considered to be gold-standard measures (e.g. the SF-36 Health Survey Questionnaire), have not been adequately evaluated in terms of responsiveness or with regard to the effect of serial measurements. Continued evaluation of existing measures is needed. Increased knowledge of the clinical and scientific properties of instruments will aid the interpretation of study results. Furthermore it will guide the refinement of existing measures, and the development of new measures to produce instruments better suited to the needs of MS rehabilitation research.

It has been suggested that rehabilitation trials should consider a profile of standardised outcome measures which could be used universally across centres to measure the various dimensions of impairment, disability, handicap and quality of life (Gladman et al 1996). This would facilitate cross-study comparisons and meta-analysis of study results.
9 Evaluation of the carers perspective

A number of surveys have highlighted the considerable assistance required by people (with moderate to severe disability) from carers in personal, domestic and social activities (McLellan et al 1989; Prouse et al 1991; Philips 1995). For example, McLellan and colleagues (1989) calculated that on average, the amount of care provided each day by the main carer was five hours and 24 minutes. Their results highlighted that this frequently impacted on both the general physical (e.g. back-ache, sleep deprivation) and emotional health (depression and anxiety) of the carers. The results of the studies undertaken in this thesis demonstrate that a short period of inpatient rehabilitation can improve the level of functional independence of patients with MS. In principle, therefore, there is good reason to expect that inpatient rehabilitation would improve the carers level and perception of burden. Furthermore, educating carers (in tasks such as transfers, bed mobility, catheterisation), and providing them with information and support are also important goals of rehabilitation (BSRM Working Party Report on MS 1993). It is therefore of interest for future studies to assess the impact of rehabilitation from a carers perspective.

10 Determining the impact of cognitive impairment on outcome

Cognitive impairment is estimated to occur in between 45-65% of patients with MS (McIntosh-Michaelis et al 1991). Specifically, deficits regularly occur on measures of recent memory, attention, information-processing speed, executive functions, and visuospatial perception (Rao 1995). It has been demonstrated that cognitive impairments play a crucial role as predictors of poor functional outcome in Stroke patients (Paolucci et al 1996). It is therefore reasonable to
expect that such impairments will impact on the outcome of rehabilitation in MS, although this has yet to be proven (Langdon and Thompson 1995). Its inclusion as a measurement variable in future studies is therefore important (Wade 1996). In particular, detailed cognitive testing may elucidate why some patients are less likely than others to maintain the gains made from rehabilitation in the longer term.

6.6 Summary

A recent editorial (Johnson 1996) expressed the view that neuroscience and neurorehabilitation can readily be linked and that there is an urgent need to foster interaction between neuroscience, clinical neurology, and neurorehabilitation at both investigative and practice levels. This viewpoint has been increasingly advocated in the past few years (Wade 1995; Thompson 1996c; Newham 1997). This philosophy is the driving force behind this thesis.

To date there is a dearth of research investigating MS inpatient rehabilitation (La Rocca and Kalb 1992). These two studies make a positive contribution to this area. By combining an experimental and clinical approach to evaluation they have made a number of methodological improvements on earlier studies. In doing so they have contributed to progressing the scientific evaluation of MS inpatient neurorehabilitation.

Firm conclusions about the effectiveness of rehabilitation cannot be drawn until a solid body of evidence has been acquired through research which is appropriate and acceptable in terms of both quality and quantity. There
remains an urgent need for further studies of comprehensive inpatient rehabilitation in people with MS, in particular to determine the duration of benefit gained and to assess and predict outcomes more accurately. The methodologies used must apply rigorous scientific analysis and review to ensure the results are scientifically credible.


292


Neurological Charities *Neurological Provision: Key Areas and Targets, Response to the 'Health of the Nation' by the Neurological Charities*. London: Department of Health, 1991.


306


Ware, J.E. and Sherbourne, C.D. The MOS 36-item short form - health survey (SF-36)


**Appendix 1 Items and scoring system of the Functional Systems Scale**

<table>
<thead>
<tr>
<th>Functional System</th>
<th>Grading Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pyramidal</strong></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>0</td>
</tr>
<tr>
<td>Abnormal signs</td>
<td>1</td>
</tr>
<tr>
<td>Minimal disability</td>
<td>2</td>
</tr>
<tr>
<td>Mild or moderate paraparesis or hemiparesis; severe monoparesis; marked paraparesis or hemiparesis; moderate quadriparesis; or monoplegia; Paraplegia, hemiplegia, or marked quadraparesis</td>
<td>3</td>
</tr>
<tr>
<td>Moderate paraparesis or hemiparesis; moderate quadriparesis; or monoplegia; Paraplegia, hemiplegia, or marked quadraparesis</td>
<td>4</td>
</tr>
<tr>
<td>Paraplegia, hemiplegia, or marked quadraparesis</td>
<td>5</td>
</tr>
<tr>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

| **Cerebellar**    |               |
| Normal            | 0             |
| Abnormal signs    | 1             |
| Mild ataxia      | 2             |
| Moderate ataxia  | 3             |
| Severe ataxia, all limbs | 4 |
| Unable to perform co-ordinated movements due to ataxia; V Unknown | 5 |
| X Used after each number when weakness (grade 3 or more on the pyramidal) interferes with testing | X |

| **Brain Stem Functions** |               |
| Normal                  | 0             |
| Signs only              | 1             |
| Moderate nystagmus or other mild disability; Severe nystagmus, marked extracranial weakness, or moderate disability of other cranial nerves; Marked dysarthria or other marked disability; Inability to swallow or speak | 2 |
| V Unknown               | V             |

| **Sensory Functions (revised 1982)** |               |
| Normal                           | 0             |
| Vibration or figure-writing decrease only, in one or two limbs | 1 |
| Mild decrease in touch or pain or position sense, and/or moderate decrease in vibration in one or two limbs; or vibratory decrease alone in three or four limbs | 2 |
| Moderate decrease in touch or pain or position sense, and/or essentially lost vibration in one or two limbs; or mild decrease in touch or pain and/or moderate decrease in all proprioceptive tests in three or four limbs | 3 |
| Marked decrease in touch or pain or loss of proprioception, alone or combined, in one or two limbs; and/or severe proprioceptive decrease in more than two limbs | 4 |
| Loss (essentially) of sensation in one or two limbs; or moderate decrease in touch or pain and/or loss of proprioception for most of the body below the head | 5 |
| Sensation essentially lost below the head; V Unknown | V |

| **Bowel and Bladder Functions (revised 1982)** |               |
| Normal                          | 0             |
| Mild urinary hesitancy, urgency, or retention | 1 |
| Moderate hesitancy, urgency, retention of bowel or bladder, or rare urinary incontinence; Frequent urinary incontinence | 2 |
| In need of almost constant catheterisation | 3 |
| Loss of bladder function | 4 |
| Loss of bladder and bowel function | 5 |
| V Unknown                      | V             |

| **Visual (or optic) Functions** |               |
| Normal                        | 0             |
| Scotoma with visual acuity (corrected) better than 20/30; Worse eye with scotoma with maximal visual acuity (corrected) of 20/30 to 20/59; Worse eye with large scotoma, or moderate decrease in fields, but with maximal visual acuity (corrected) of 20/60 to 20/99; Worse eye with marked decrease of fields and maximal visual acuity (corrected) of 20/100 to 20/200; grade 3 plus maximal acuity of better eye of 20/60 or less; Worse eye with maximal visual acuity (corrected) less than 20/200; grade 4 plus maximal acuity of better eye of 20/60 or less | 1 |
| Grade 5 plus maximal acuity of better eye of 20/60 or less; V Unknown | 2 |
| X Added to grades 0 to 6 for presence of temporal pallor | X |

| **Cerebral (or mental) Functions** |               |
| Normal                        | 0             |
| Mood alteration only; Mild decrease in mentation; Moderate decrease in mentation; Marked decrease in mentation (chronic brain syndrome - moderate); Dementia or chronic brain syndrome-severe or incompetent | 1 |
| V Unknown                      | V             |

| **Other Functions** |               |
| 1.None; 2 any other neurological findings attributed to MS (specify); V Unknown | 1 |

---

The Functional Systems Scale provides a detailed grading system for various neurological functions, allowing for a comprehensive assessment of a patient's condition. The scale is designed to reflect the severity and type of neurological impairments, ensuring that each aspect of function is accurately scored. This system is particularly useful in tracking the progression of conditions such as multiple sclerosis (MS), where understanding the impact on different systems is crucial for effective management and care.
### Appendix 2 Items and scoring system of the Expanded Disability Status Scale.

<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Normal neurological exam (all FS grades 0; Cerebral grade 1 acceptable).</td>
</tr>
<tr>
<td>1.0</td>
<td>No disability, minimal signs in one FS (i.e. grade 1 excluding cerebral grade 1).</td>
</tr>
<tr>
<td>1.5</td>
<td>No disability, minimal signs in more than one FS (more than one grade 1 excluding Cerebral grade1).</td>
</tr>
<tr>
<td>2.0</td>
<td>Minimal disability in one FS (one FS grade 2, others 0 or 1).</td>
</tr>
<tr>
<td>2.5</td>
<td>Minimal disability in two FS (two FS grade 2, others 0 or 1).</td>
</tr>
<tr>
<td>3.0</td>
<td>Moderate disability in one FS (one FS grade 3, others 0 or 1), or mild disability in three or four FS (three/four FS grade 2, others 0 or 1) though fully ambulatory.</td>
</tr>
<tr>
<td>3.5</td>
<td>Fully ambulatory but with moderate disability in one FS (one grade 3) and one or two FS grade 2; or two FS grade 3; or five FS grade 2 (others 0 or 1).</td>
</tr>
<tr>
<td>4.0</td>
<td>Fully ambulatory without aid, self-sufficient, up and about some 12 hours a day despite relatively severe disability consisting of one FS grade 4 (others), or combinations of lesser grades exceeding limits of previous steps. Able to walk without aid or rest some 500 meters.</td>
</tr>
<tr>
<td>4.5</td>
<td>Fully ambulatory without aid, up and about much of the day, able to work a full day, may otherwise have some limitation of full activity or require minimal assistance; characterised by relatively severe disability, usually consisting of one FS grade 4 (others 0 or 1) or combinations of lesser grades exceeding the limits of previous steps. Able to walk without aid or rest for some 300 meters.</td>
</tr>
<tr>
<td>5.0</td>
<td>Ambulatory without aid or rest for about 200 meters; disability severe enough to impair full daily activities (e.g. to work full day without special provisions). (Usual FS equivalents are one grade 5 alone, others 0 or 1; or combinations of lesser grades usually exceeding specifications for step 4.)</td>
</tr>
<tr>
<td>5.5</td>
<td>Ambulatory without aid or rest for about 100 meters; disability severe enough to preclude full daily activities. (Usual FS equivalents are grade 5 alone, others 0 or 1; or combinations of lesser grades usually exceeding those for step 4).</td>
</tr>
<tr>
<td>6.0</td>
<td>Intermittent or unilateral constant assistance (cane, crutch, or brace) required to walk about 100 meters with or without resting. (Usual FS equivalents are combinations with more than two FS grade 3+.)</td>
</tr>
<tr>
<td>6.5</td>
<td>Constant bilateral assistance (canes, crutches, or braces) required to walk about 100 metres without resting. (Usual FS equivalents are combinations with more than two FS grade 3+).</td>
</tr>
<tr>
<td>7.0</td>
<td>Unable to walk beyond about 5 metres even with aid, essentially restricted to wheelchair; wheels self in standard wheelchair and transfers alone; up and about in w/c some 12 hours a day. (Usual FS equivalents are combinations with more than one FS grade 4+; very rarely, pyramidal grade 5 alone)</td>
</tr>
<tr>
<td>7.5</td>
<td>Unable to take more than a few steps; restricted to wheelchair; may need aid in transfer; wheels self but cannot carry on in standard wheelchair a full day; may require motorised wheelchair. (Usual FS equivalents are combinations with more than one FS grade 4+)</td>
</tr>
<tr>
<td>8.0</td>
<td>Essentially restricted to bed or chair or perambulated in wheelchair, but may be out of bed itself much of the day; retains many self care functions; generally has effective use of arms. (Usual FS equivalents are combinations, generally grade 4+, in several systems).</td>
</tr>
<tr>
<td>8.5</td>
<td>Essentially restricted to bed much of the day; has some effective use of arm(s); retains some self-care functions. (Usual FS equivalents are combinations, generally grade 4+ in several systems)</td>
</tr>
<tr>
<td>9.0</td>
<td>Helpless bed patient; can communicate and eat (Usual FS equivalents are combinations, mostly grade 4+)</td>
</tr>
<tr>
<td>9.5</td>
<td>Totally helpless bed patient; unable to communicate effectively or eat/swallow. (Usual FS equivalents are combinations, almost all grade 4+).</td>
</tr>
<tr>
<td>10.0</td>
<td>Death due to MS</td>
</tr>
</tbody>
</table>
Appendix 3  Items and scoring system of the Functional Independence Measure (UDS Version 4.0, 1994)

a) Items

<table>
<thead>
<tr>
<th>Motor Domain</th>
<th>Cognitive Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self Care</td>
<td>Communication</td>
</tr>
<tr>
<td>Feeding</td>
<td>Comprehension</td>
</tr>
<tr>
<td>Grooming</td>
<td>Expression</td>
</tr>
<tr>
<td>Bathing</td>
<td>Social Cognition</td>
</tr>
<tr>
<td>Dressing-upper body</td>
<td>Social Interaction</td>
</tr>
<tr>
<td>Dressing-lower body</td>
<td>Problem Solving</td>
</tr>
<tr>
<td>Toileting</td>
<td>Memory</td>
</tr>
<tr>
<td>Sphincter Control</td>
<td></td>
</tr>
<tr>
<td>Bladder management</td>
<td></td>
</tr>
<tr>
<td>Bowel management</td>
<td></td>
</tr>
<tr>
<td>Transfers</td>
<td></td>
</tr>
<tr>
<td>Bed, Chair, Wheelchair</td>
<td></td>
</tr>
<tr>
<td>Toilet</td>
<td></td>
</tr>
<tr>
<td>Bath, Shower</td>
<td></td>
</tr>
<tr>
<td>Locomotion</td>
<td></td>
</tr>
<tr>
<td>Walk/Wheelchair</td>
<td></td>
</tr>
<tr>
<td>Stairs</td>
<td></td>
</tr>
</tbody>
</table>

b) Scoring system

<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>Independent (no helper) Complete independence: all tasks performed safely,</td>
</tr>
<tr>
<td></td>
<td>without modification, or aids, and within a reasonable amount of time.</td>
</tr>
<tr>
<td>6</td>
<td>Modified independence: requiring an aid, safety considerations exist, or</td>
</tr>
<tr>
<td></td>
<td>takes longer than a reasonable time.</td>
</tr>
<tr>
<td>5</td>
<td>Dependent (requires helper) Supervision or set-up: standby, cueing, coaxing</td>
</tr>
<tr>
<td></td>
<td>or help with set-up is required.</td>
</tr>
<tr>
<td>4</td>
<td>Minimal contact assistance: minimal assistance required, subject expends</td>
</tr>
<tr>
<td></td>
<td>75% or more of the effort.</td>
</tr>
<tr>
<td>3</td>
<td>Moderate assistance: subject expends between 50-75% of the effort.</td>
</tr>
<tr>
<td>2</td>
<td>Maximal assistance: subject expends between 25-50% of the effort.</td>
</tr>
<tr>
<td>1</td>
<td>Total assistance: subject expends less than 25% of the effort.</td>
</tr>
</tbody>
</table>
Appendix 4 Items of the London Handicap Scale

1. Does your health stop you from getting around?
(Think about how you get from one place to another, using any help, aids or means of transport that you normally have available)

<table>
<thead>
<tr>
<th>Degree of Disability</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not at all</td>
<td>You go everywhere you want, no matter how far away</td>
</tr>
<tr>
<td>Very slightly</td>
<td>You go most places you want, but not all</td>
</tr>
<tr>
<td>Quite a lot</td>
<td>You get out of the house, but not far away from it</td>
</tr>
<tr>
<td>Very much</td>
<td>You don't go outside, but you can move around from room to room indoors</td>
</tr>
<tr>
<td>Almost completely</td>
<td>You are confined to a single room, but can move around in it</td>
</tr>
<tr>
<td>Completely</td>
<td>You are confined to a bed or a chair. You cannot move around at all. There is no-one to move you.</td>
</tr>
</tbody>
</table>

2. Does your health stop you looking after yourself?
(Think about things like housework, shopping, looking after money, cooking, laundry, getting dressed, washing, shaving and using the toilet)

<table>
<thead>
<tr>
<th>Degree of Disability</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not at all</td>
<td>You can do everything to look after yourself</td>
</tr>
<tr>
<td>Very slightly</td>
<td>You need a little help now and again</td>
</tr>
<tr>
<td>Quite a lot</td>
<td>You need help with some tasks (such as heavy housework or shopping), but no more than once a day</td>
</tr>
<tr>
<td>Very much</td>
<td>You do some things for yourself, but you need help more than once a day. You can be left alone safely for a few hours.</td>
</tr>
<tr>
<td>Almost completely</td>
<td>You need help to be available all the time. You cannot be left alone safely.</td>
</tr>
<tr>
<td>Completely</td>
<td>You need help with everything. You need constant attention, day and night</td>
</tr>
</tbody>
</table>

3. Does your health limit your work or leisure activities?
(Think about things like work (paid or not), housework, gardening, sports, hobbies, going out with friends, travelling, reading, looking after children, watching television and going on holiday)

<table>
<thead>
<tr>
<th>Degree of Disability</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not at all</td>
<td>You do everything you want to do</td>
</tr>
<tr>
<td>Very slightly</td>
<td>You do almost all the things you want to do</td>
</tr>
<tr>
<td>Quite a lot</td>
<td>You find something to do almost all the time, but cannot do some things for as long as you would like</td>
</tr>
<tr>
<td>Very much</td>
<td>You are unable to do a lot of things, but can find something to do most of the time</td>
</tr>
<tr>
<td>Almost completely</td>
<td>You are unable to do most things, but can find something to do some of the time</td>
</tr>
<tr>
<td>Completely</td>
<td>You sit all day doing nothing. You cannot keep yourself busy or take part in any activities</td>
</tr>
</tbody>
</table>

continued over page

318
### Appendix 4 continued. Items of the London Handicap Scale

#### 4. Does your health stop you from getting on with people?
*(Think about family, friends and the people you might meet during a normal day)*

<table>
<thead>
<tr>
<th>Rating</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not at all</td>
<td>You get on well with people, see everyone you want to see, and meet new people</td>
</tr>
<tr>
<td>Very slightly</td>
<td>You get on well with people, but your social life is slightly limited</td>
</tr>
<tr>
<td>Quite a lot</td>
<td>You are fine with people you know well, but you feel uncomfortable with strangers</td>
</tr>
<tr>
<td>Very much</td>
<td>You are fine with people you know well, but you have few friends and little contact with neighbours. Dealing with strangers is very hard.</td>
</tr>
<tr>
<td>Almost completely</td>
<td>Apart from the people who look after you, you see no-one. You have no friends and no visitors</td>
</tr>
<tr>
<td>Completely</td>
<td>You don’t get on with anyone, not even the people who look after you.</td>
</tr>
</tbody>
</table>

#### 5. Does your health stop you understanding the world around you?
*(Think understanding the world about you, and finding your way around in it)*

<table>
<thead>
<tr>
<th>Rating</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not at all</td>
<td>You fully understand the world around you. You see, hear, speak and think clearly, and your memory is good.</td>
</tr>
<tr>
<td>Very slightly</td>
<td>You have problems with hearing, speaking, seeing or your memory, but these do not stop you doing most things</td>
</tr>
<tr>
<td>Quite a lot</td>
<td>You have problems with hearing, speaking, seeing or your memory which make life difficult a lot of the time. But you understand what is going on</td>
</tr>
<tr>
<td>Very much</td>
<td>You have (he/she has) great difficulty understanding what is going on</td>
</tr>
<tr>
<td>Almost completely</td>
<td>He/she is unable to tell where he/she is or what day it is. He/she cannot look after him/herself at all</td>
</tr>
<tr>
<td>Completely</td>
<td>He/she is unconscious, completely unaware of anything going on around him/her</td>
</tr>
</tbody>
</table>

#### 6. Are you able to afford the things you need?
*(Think about whether health problems have led to any extra expenses, or have caused you to earn less than you would if you were healthy)*

<table>
<thead>
<tr>
<th>Rating</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes, easily</td>
<td>You can afford everything you need. You have easily enough money to buy modern labour-saving devices, and anything you may need because of ill-health</td>
</tr>
<tr>
<td>Fairly easily</td>
<td>You have just about enough money. It is fairly easy to cope with expenses caused by ill-health</td>
</tr>
<tr>
<td>Just about</td>
<td>You are less well off than other people like you; however, with sacrifices you can get by without help</td>
</tr>
<tr>
<td>Not really</td>
<td>You only have enough money to meet your basic needs. You are dependent on state benefits for any extra expenses you have because of ill-health</td>
</tr>
<tr>
<td>No</td>
<td>You are dependent on state benefits, or money from other people or charities. You cannot afford things you need</td>
</tr>
<tr>
<td>Absolutely not</td>
<td>You have no money at all and no state benefits. You are totally dependent on charity for your most basic needs</td>
</tr>
</tbody>
</table>
Appendix 5 Items of the 36 item Short Form Health Survey Questionnaire

1. In general, would you say your health is:
   Excellent  Very good  Good  Fair  Poor

2. Compared to one year ago how would you rate your health in general now?
   Much better  Somewhat better  About the same  Somewhat worse  Much worse

3. The following questions are about activities you might do during a typical day. Does your health now limit you in these activities? If so, how much?
   a) Vigorous activities, such as running, lifting heavy objects, participating in strenuous sports
   b) Moderate activities, such as moving a table, pushing a vacuum cleaner, bowling or playing golf
   c) Lifting or carrying groceries
   d) Climbing several flights of stairs
   e) Climbing one flight of stairs
   f) Bending, kneeling or stooping
   g) Walking more than a mile
   h) Walking half a mile
   i) Walking one hundred yards
   j) Bathing or dressing yourself
   Yes limited a lot  Yes, limited a little  No, not limited at all

4. During the past 4 weeks have you had any of the following problems with your work or other regular daily activities as a result of your physical health?
   a) Cut down on the amount of time you spent on work or other activities
   b) Accomplished less than you would like to
   c) Were limited in the kind of work or other activities
   d) Had difficulty performing the work or other activities (for example, it took extra effort)
   Yes  No

5. During the past 4 weeks have you had any of the following problems with your work or regular daily activities as a result of emotional problems (such as feeling depressed or anxious)?
   a) Cut down on the amount of time you spent on work or other activities
   b) Accomplished less than you would like to
   c) Didn't do work or other activities as carefully as usual
   Yes  No

6. During the past 4 weeks, to what extent has your physical health or emotional problems interfered with your normal social activities with family, friends, neighbours or groups?
   Not at all  Slightly  Moderately  Quite a bit  Extremely

7. How much bodily pain have you had during the past 4 weeks?
   None  Very mild  Mild  Moderate  Severe  Very severe

continued over page........................
Appendix 5 continued. Items of the Short form Health Survey Questionnaire

8. During the past 4 weeks, how much did pain interfere with your normal work (including booth work outside the home and housework)?
   Not at all  A little bit  Moderately  Quite a bit  Extremely

9. These questions ask you about how you feel and how things have been with you during the past 4 weeks. For each question, please give the one answer that comes closest to the way you have been feeling. How much of the time during the past 4 weeks:
   a) did you feel full of life?
   b) have you been a very nervous person?
   c) have you felt so down in the dumps that nothing could cheer you up?
   d) have you felt calm and peaceful?
   e) did you have a lot of energy?
   f) have you felt downhearted and low?
   g) did you feel worn out?
   h) have you been a happy person?
   i) did you feel tired?
   All of the time  Most of the time  A good bit of the time  Some of the time
   A little of the time  None of the time

10. During the past 4 weeks, how much of the time has your physical health or emotional problems interfered with your social activities (like visiting with friends, relatives etc.)?
    All of the time  Most of the time  A good bit of the time  Some of the time
    A little of the time  None of the time

11. How TRUE or FALSE is each of the following statements for you?
    a) I seem to get ill a little more easily than other people
    b) I am as healthy as anybody I know
    c) I expect my health to get worse
    d) My health is excellent
    Definitely true  Mostly true  Don't know  Mostly false  Definitely false

(UK English Language Version)
Appendix 6: Items of the 28 item General Health Questionnaire (GHQ-28)

### HAVE YOU RECENTLY (in the past few weeks):

1. **Been feeling perfectly well and in good health?**
   - better than usual
   - same as usual
   - less than usual
   - much less than usual

2. **Been feeling in need of a good tonic?**
   - not at all
   - no more than usual
   - rather more than usual
   - much more than usual

3. **Been feeling run down and out of sorts?**
   - not at all
   - no more than usual
   - rather more than usual
   - much more than usual

4. **Felt that you are ill?**
   - not at all
   - no more than usual
   - rather more than usual
   - much more than usual

5. **Been getting any pains in your head?**
   - not at all
   - no more than usual
   - rather more than usual
   - much more than usual

6. **Been getting a feeling of tightness or pressure in your head?**
   - not at all
   - no more than usual
   - rather more than usual
   - much more than usual

7. **Been having hot or cold spells?**
   - not at all
   - no more than usual
   - rather more than usual
   - much more than usual

8. **Lost much sleep over worry?**
   - not at all
   - no more than usual
   - rather more than usual
   - much more than usual

9. **Had difficulty staying asleep once you are off?**
   - not at all
   - no more than usual
   - rather more than usual
   - much more than usual

10. **Felt constantly under strain?**
    - not at all
    - no more than usual
    - rather more than usual
    - much more than usual

11. ** Been getting edgy and bad tempered?**
    - not at all
    - no more than usual
    - rather more than usual
    - much more than usual

12. **Been getting scared or panicky for no good reason?**
    - not at all
    - no more than usual
    - rather more than usual
    - much more than usual

13. **Found everything getting on top of you?**
    - not at all
    - no more than usual
    - rather more than usual
    - much more than usual

14. **Been feeling nervous and strung up all the time?**
    - not at all
    - no more than usual
    - rather more than usual
    - much more than usual

15. **Been managing to keep yourself busy and occupied?**
    - more so than usual
    - same as usual
    - rather less than usual
    - much less than usual

16. **Been taking longer over the things you do?**
    - quicker than usual
    - same as usual
    - longer than usual
    - much longer than usual

17. **Felt on the whole you were doing things well?**
    - better than usual
    - about the same
    - less well than usual
    - much less well

18. **Been satisfied with the way you've carried out the task?**
    - more satisfied
    - about same as usual
    - less satisfied than usual
    - much less satisfied

19. **Felt that you are playing a useful part in things?**
    - more so than usual
    - same as usual
    - less useful than usual
    - much less useful

20. **Felt capable of making decisions about things?**
    - more so than usual
    - same as usual
    - less so than usual
    - much less capable

21. **Been able to enjoy your normal day-to-day activities?**
    - more so than usual
    - same as usual
    - less so than usual
    - much less capable

22. **Been thinking of yourself as a worthless person?**
    - not at all
    - no more than usual
    - rather more than usual
    - much more than usual

23. **Felt that life is entirely hopeless?**
    - not at all
    - no more than usual
    - rather more than usual
    - much more than usual

24. **Felt that life isn't worth living?**
    - not at all
    - no more than usual
    - rather more than usual
    - much more than usual

25. **Thought of the possibility of making away with yourself?**
    - not at all
    - no more than usual
    - rather more than usual
    - much more than usual

26. **Found at times you couldn't do anything because your nerves were too bad?**
    - not at all
    - no more than usual
    - rather more than usual
    - much more than usual

27. **Found yourself wishing you were dead and away from it all?**
    - not at all
    - no more than usual
    - rather more than usual
    - much more than usual

28. **Found that the idea of taking your own life kept coming into your mind?**
    - definitely not
    - I don't think so
    - has crossed my mind
    - definitely has
A study to investigate the benefits of rehabilitation in people with MS

WHAT ARE WE DOING?
At the Neurorehabilitation Unit at Finchley we are conducting a study to look at the benefits of rehabilitation on people, such as yourself, with multiple sclerosis. The study hopes to involve almost every person with multiple sclerosis who is admitted to this unit over the next twelve months.

WHY ARE WE DOING THIS STUDY?
It is generally agreed that rehabilitation is important for people with multiple sclerosis. This has yet to be proven. In order to do this we must investigate whether benefits occur as a result of rehabilitation. To date very little work has been undertaken to evaluate this. The aim of this study is to demonstrate which changes occur and to describe them.

HOW ARE YOU INVOLVED IN THIS?
Today you will be assessed by Dr Thompson and a team of therapy and nursing staff to see if they think you may benefit from admission to the rehabilitation unit at Finchley. If admission to the unit is recommended we might ask you to be involved in this study.

If you agree to being involved you will be randomly allocated into one of two groups:
• People in Group A will be booked to begin their rehabilitation programme at Finchley in 6 weeks time.
• People in Group B will be booked to begin their rehabilitation programme as soon as the waiting list permits.

WILL THIS AFFECT THE REHABILITATION PROGRAMME YOU PARTICIPATE IN?
NO. The programme you will participate in will NOT be affected by your involvement in this study. As is normal practice for this unit, you will have a programme tailored specifically for your individual needs. Participation in this study will NOT affect your normal medical care.

CAN YOU WITHDRAW FROM THE STUDY IF YOU WISH?
YES you are free to withdraw from the study at any time. This will NOT affect your normal medical care.

WHAT HAPPENS NOW?
If you agree to participate in this study, you will be assessed by a Research Doctor and Therapist and asked questions regarding the difficulties you may experience as a result of multiple sclerosis.
• If you are in group A these assessments will take place today and on admission to the Unit.
• If you are in Group B these assessments will take place on admission to the Unit and six weeks following.

DO YOU HAVE ANY QUESTIONS?
If you have any questions, or would like any aspects to be described in more detail, please do not hesitate to speak to Dr Thompson or Jenny Freeman either today or by telephone. We are very happy to answer any of your questions.
Appendix 8  Form for demographic and diagnostic details on admission

Date: ______  Patient Identification Number: ______

Assessment Number: [Clinic] [Admit] [6 week] [D/C] [3mth] [6mth] [9mth] [12mth]

Information gathered from: 1. Patient  2. Carer

Born and educated in Britain?  1. Yes  2. No

Sex:  1. Male  2. Female  Age (in completed years): ______

Diagnosis:  1. Primary Progressive  2. Secondary Progressive

First symptoms of MS: (month) __ (year) __  Diagnosis of MS: (month) __ (year) __

Symptom progression over past 6 months: 1. Stable  2. Progressive

Recent Relapse: 1. Yes  2. No  If yes, how many months ago was last relapse: __

Any health problems, other than MS: 1. Yes  2. No

If Yes, please specify: ____________________________________________


Type of housing: 1. Owner/occupied  2. Rented from local authority


Is housing suitable: 1. Yes  2. No

Any alterations/additions made to home because of disability: 1. Yes  2. No

If yes, please specify: ____________________________________________


3. Lives with friends/carer  4. Other

Main carer: 1. No carer assistance required  2. Spouse  3. Relative

4. Friend/Neighbour  5. Paid Carer  6. Other

Is main carer able to look after you: 1. No  2. Yes, with formal services

3. Yes, without formal services  3. Not applicable


Appendix 9  FIM intrarater reliability study

Aim: to determine the intrarater reliability of the researcher in scoring the FIM motor domain subscales and total score.

Sample: 25 MS patients undertaking inpatient rehabilitation, representing a wide range of types and severity of disabilities.

Procedure: The FIM scores were derived from patient interview by the researcher. Two assessments were undertaken three days apart. The researcher did not have further access to the initial scores. In accordance with the guidelines, scores were determined by actual performance of tasks on a daily basis, rather than optimal performance.

Analysis: The intraclass correlation coefficient (ICC) was used to determine intrarater reliability. This statistic measures the ratio of between-subject variance to the total variance which includes both between-subject and within-subject variance (Guyatt et al 1992).

Results:

<table>
<thead>
<tr>
<th>FIM subscales / domain</th>
<th>ICC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selfcare subscale</td>
<td>0.99</td>
</tr>
<tr>
<td>Sphincter subscale</td>
<td>1.00</td>
</tr>
<tr>
<td>Transfers subscale</td>
<td>0.99</td>
</tr>
<tr>
<td>Locomotion subscale</td>
<td>0.99</td>
</tr>
<tr>
<td>Total FIM motor domain</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Conclusion:
The FIM ratings undertaken by the researcher are highly reliable both in terms of the subscale scores and the total motor domain scores.
Appendix 10
Determining the validity of the self-report method for scoring the FIM

Aim: The self-report method has been previously shown to be a valid and reliable method of scoring the FIM in MS patients (Brousseau and Wolfson 1994). The aim of this study was to determine whether the self-report method of scoring the FIM was valid in this population of MS rehabilitation inpatients.

Sample: Fifty rehabilitation inpatients with progressive MS who were participating in either the randomised controlled study or the longitudinal study.

Procedure: Multidisciplinary team ratings based on observation is the recommended method of scoring the FIM. Scores determined by this method were routinely recorded for all patients within the first week of admission to the unit. Members of the team were blinded to the group allocation of the patient. The researcher scored the FIM by patient interview within 24 hours of admission to the unit.

Analysis: Multidisciplinary team scores were correlated with the self-report scores using Spearmann's Correlation Coefficient.

Results: The scatterplots and correlation coefficient values show the strong relationship between the team and the self-report FIM motor domain scores, on both admission and discharge.

\[
\text{Admission: } r = 0.95 \\
\text{Discharge: } r = 0.96
\]

Conclusion: These results support the self-report interview as a valid method of rating the FIM motor domain in this population.
Appendix 11 Information leaflet for patients in the longitudinal trial

A study to investigate the benefits of rehabilitation in people with MS

WHAT ARE WE DOING?
At the Neurorehabilitation Unit at Finchley we are conducting a study to look at the benefits of rehabilitation on people, such as yourself, with multiple sclerosis.
The study hopes to involve almost every person with multiple sclerosis who is admitted to this unit over the next twelve months.

WHY ARE WE DOING THIS STUDY?
It is generally agreed that rehabilitation is important for people with multiple sclerosis. This has yet to be proven. In order to do this we must investigate whether benefits occur as a result of rehabilitation.
To date very little work has been undertaken to evaluate this.
The aim of this study is to demonstrate which changes occur and to describe what happens to these changes over the year following discharge.

HOW ARE YOU INVOLVED IN THIS?
Today you will be assessed by Dr Thompson and a team of therapy and nursing staff to see if they think you may benefit from admission to the rehabilitation unit at Finchley. If admission to the unit is recommended we might ask you to be involved in this study.
If you agree to participate then on the first day of your admission to the rehabilitation unit you will be assessed by a research Doctor and Therapist and asked to fill out a series of questionnaires. These assessments will take place on admission and on discharge from the unit, and then at three monthly intervals for the following year.

WILL THIS AFFECT THE REHABILITATION PROGRAMME YOU PARTICIPATE IN?
NO. The programme you will participate in will NOT be affected by your involvement in this study. As is normal practice for this unit, you will have a programme tailored specifically for your individual needs. Participation in this study will NOT affect your normal medical care.

CAN YOU WITHDRAW FROM THE STUDY IF YOU WISH?
YES you are free to withdraw from the study at any time. This will NOT affect your normal medical care.

WHAT HAPPENS NOW?
If you agree to participate in this study, you will be assessed by a Research Doctor and Therapist and asked questions regarding the difficulties you may experience as a result of multiple sclerosis. These assessments and questionnaires will be undertaken on the first day of your admission to the rehabilitation unit, on discharge from the unit, and then at three monthly intervals for the following year in the outpatient clinic at the National Hospital for Neurology and Neurosurgery, Queen Square.

DO YOU HAVE ANY QUESTIONS?
If you have any questions, or would like any aspects to be described in more detail, please do not hesitate to speak to Dr Thompson or Jenny Freeman either today or by telephone. We are very happy to answer any of your questions.
Appendix 12 Form for demographic and diagnostic details on discharge

| Date: __________ | Patient Identification Number : __ _ |
| Assessment Number: [Clinic] [Admit] [6 weeks] [D/C] [3mths] [6mths] [9mths] [12mth] |
| Information gathered from: 1. Patient 2. Carer |
| Date admitted: ______ | Date discharged: ______ | Length of Stay: (days): ______ |
| Admitted from: 1. Home 2. Another Rehabilitation Unit 3. Hospital |
| 4. Other ________ |
| Discharged to: 1. Home 2. Another Rehabilitation Unit 3. Hospital |
| 4. Other ________ |

**Current Services:**

**Additional services recommended from the NRU:**

**Current Outpatient Therapy:**

**Additional therapy services recommended:**

**Recommended for re-housing:** 1. Yes 2. No

**Recommended for house adaptations:** 1. Yes 2. No

List Recommendations for house adaptations: ___________________________
Appendix 13 Form for demographic and diagnostic details on follow-up

Date: ___________ Patient Identification Number: __ __

Follow-up Assessment Number: [3mths] [6mths] [9mths] [12mths]

Information gathered from: 1. Patient 2. Carer

Have your home arrangements changed recently: 1 Yes 2. No

If yes then please fill out another initial demographic sheet.

Relapse since last review? 1. Yes 2 No If yes, when?_____________

Re-hospitalisation since last review? 1. Yes 2. No

If yes, why? ___________________________________________________

Recommended Services:

Refer to d/c demographic sheet, ring any that apply and code as:

1 = received + ongoing ; 2 = received + completed; 3 = not received

9. Personal Care Attendant 10. Other___________

Recommended O/P Therapy:

Refer to d/c demographic sheet, ring any that apply and code as:

1 = received + ongoing ; 2 = received + completed; 3 = not received

8. Orthotic Services 9. Psychology/Counselling 10. Other, __________

If recommended for re-housing then is this currently:

1. underway 2. completed 3. not commenced

If recommended for house adaptations then are they currently:

1. underway 2. completed 3. not commenced
Appendix 14 Profile of patients who experienced a relatively rapid reduction in functional ability between two consecutive follow-up assessments.

Case 4: deteriorated by 15 FIM points between discharge and three months. This coincided with:
- hospitalisation for management of severe pain and spasms
- a deterioration in overall neurological severity (EDSS score increased by one point from 8.0 to 9.0).

A further deterioration of 17 FIM points occurred between the three month and six month assessment. During this time Mr N remained hospitalised for continued management of the intractable spasms and pain. He died of pneumonia prior to the 12 month assessment

Case 6: deteriorated by 19 FIM points between discharge and three months. This coincided with:
- a deterioration by two points in the Visual System of the FS Scale.

It was also notable that the home adaptations recommended on discharge were not undertaken as Mrs A decided to move to a bungalow. At the six month assessment, after Mrs A had moved into her new home, the FIM scores had increased to five points above the initial assessment score.

Case 16: deteriorated by 16 FIM points between three months and six months. This coincided with:
- hospitalisation for surgical intervention (colostomy and insertion of a suprapubic catheter).
- a deterioration in overall neurological severity: EDSS score increased by one point from 8.0 to 9.0; both the Visual and Cerebellar Systems of the FS Scale increased by two points.

Case 36: deteriorated by 17 FIM points between three and six months. This coincided with:
- deterioration in overall neurological severity: EDSS scores increased by 0.5 point from 6.5 to 7.0 (i.e. walking to essentially wheelchair bound), and by one point on the Pyramidal, Brainstem and Cerebral Systems of the FS Scale.
- a diagnosis of depression by the neurologist
- moving into different accommodation