The assessment of posture and balance post-stroke

A thesis submitted for the degree of doctor of philosophy

by

Sarah F Tyson

ABSTRACT

Physiotherapy for people with stroke has been found to be beneficial but details of the most effective interventions are unclear. Further development of the evidence base for stroke physiotherapy is limited by a lack of clinical practice models, sensitive clinically based outcome measures and effective stratification techniques to characterise homogenous groups of subjects. These issues are addressed here with regard to balance and posture. These aspects were chosen because they form a cornerstone of stroke physiotherapy as they are thought essential for the rehabilitation of functional activities.

A systematic review of assessment methods in the literature revealed a lack of measurement tools which met the utility criteria: reliability, validity, sensitivity to short-term change, suitability for a wide range of abilities, ease of use and suitability for different settings. This prompted the development of a new measurement tool. Firstly, a model of the clinical assessment process was developed using an adapted focus group method with neurological physiotherapists. This informed the content of a new measurement tool which combined an ordinal scale with functional performance tests- the Brunel Balance Assessment. The tool was evaluated in a series of studies involving 92 stroke patients. It was hierarchical (coefficient of reproducibility= 0.99, coefficient of scalability = 0.69), reliable (100% agreement) and valid as a measure of balance disability (r=0.58-0.97). The psychometric properties of the individual functional performance tests were also tested and found to be reliable (ICCs =0.88-1) and valid (r=0.32-0.63). Measurement error ranged 0-40% and the minimum change needed to detect true clinical change was calculated for each test.

Balance disability, measured with the Brunel Balance Assessment, is heterogeneous with sitting, standing and stepping balance forming distinct levels of ability ($p<0.027$). Consequently, the BBA could be used to stratify people with stroke according to balance ability. Weakness, sensation and age were significant independent contributors to balance disability (r^2=82.7%). Balance ability was a strong contributor to independence in ADL ($p<0.0001$).

The findings of this thesis address the issues that have limited research into stroke physiotherapy with regard to balance disability. In relation to clinical practice, a robust measurement and stratification tool has been developed.

Table of Contents

List of Tables

List of Figures

Appendices

ACKNOWLEDGEMENTS

I wish to acknowledge and thank the many people who have supported and encouraged me during the preparation of this thesis.

First and foremost is my supervisor and head of department - Prof. Lorraine DeSouza, whose wise advice, support and encouragement has kept me going through thick and thin. I would also like to thank my friends and colleagues in the Dept of Health Studies - Sandra Naylor, Janet Lewis, Betty Bond for their moral and practical support, Conor Gissane and Lousie O'Hara for their statistical advice. Dr Agnes Shiel of the MRC Cognition and Brain Science Unit in Cambridge for her mighty red pen. I am indebted and very grateful to Prof Simon Millar of the University of Newcastle and Mr Andy Vail of the University of Manchester for their invaluable and most generous expert advice. My thanks also go to the people with stroke and the neurological physiotherapists in the participating hospitals for their time and enthusiasm.

The Dept of Health Studies at Brunel University and the Brunel University Research Enterprise and Initiative Fund provided funding and this is gratefully acknowledged.

Finally, the contribution of my husband, David Cotton and my sons Robbie and Alex is acknowledged with love and gratitude for putting up with it all.

This thesis is dedicated to the memory of my maternal grandmother, Janet Bunkle for teaching me that most things are possible.

Chapter 1

Introduction

1.1 Introduction

Stroke is the most common cause of adult disability in Britain (Martin 1989). Every year some 100,000 new strokes occur (Bamford et al 1988), at any one time up to 20% of acute medical beds are occupied by people with stroke (Wade 1994), and up to 4% of the NHS budget is spent on cerebro-vascular disease and its aftermath (Office of Health Economics 1988). The personal consequences of stroke can be devastating. Although 80% of people with stroke survive the initial insult (Bamford et al 1990) and most make some spontaneous recovery, some 50% of survivors will have a significant long-term disability due to motor, communication and/or cognitive impairments, while approximately 12% will require institutionalised care (Legh-Smith et al 1986).

Rehabilitation aims to tackle the consequences of stroke. Definitions of rehabilitation vary but the one used by the Effective Health Bulletin on Stroke Rehabilitation (1992) is typical in stating that rehabilitation aims to aid physical recovery, encourage independence, promote adaptation to stroke-related disability and handicap, and prevent secondary complications. There is now strong evidence that stroke rehabilitation is effective (Stroke Unit Stroke Unit Trialists Collaborations 1998). When delivered by a co-ordinated, multi-disciplinary specialist staff working from a geographically identified base, rehabilitation can reduce, mortality and morbidity, improve the degree and speed of recovery, reduce length of stay in hospital and affect destination on discharge (Kalra 1994; Indredavik et al 1997; Stroke Unit Stroke Unit Trialists Collaborations' Collaboration 1998; Ronning & Guldvog 1998a & b).

Details of which aspects of this process of care deliver these improvements are uncertain. There is evidence that good outcome follows effective multi-disciplinary team-working (Stroke Unit Stroke Unit Trialists Collaborations' Collaboration 1998; Royal College of Physicians 2000), the use of goal-setting (Royal College of Physicians 2000: Wade 1998a &1999), staff education, functional training, and integrated physiotherapy and nursing (Indredavik et al 1997). But the most effective aspects of therapy are uncertain (Royal College of Physicians 2000).

The intensity or amount of therapy influences outcome, and the more therapy the better the outcome (Smith et al 1981; Rappaport & Eerd 1989; Langhorne et al 1996; Kwakkel et al 1997; 1999) although it is not clear how much therapy is optimal and whether there are upper or lower thresholds.

The timing of therapy may also be a factor. It is difficult to distinguish the effect of early physiotherapy intervention from the effects of well-organised stroke care, and/or spontaneous recovery, but several reviews have concluded that early physiotherapy intervention achieves better outcomes (Wagenaar & Meijer 1991 a&b; Ashburn et al 1993; Cifu &Stewart 1999). Indredavik et al (1997) found that early mobilisation and functional training (which is the domain of physiotherapy) improved outcome following acute stroke. There is stronger evidence that the effectiveness of physiotherapy is not limited to the immediate post-acute recovery period (three-six months), and that physiotherapy can promote functional improvement long after stroke (Wade et al 1992; Dean & Shepherd 1997; Royal College of Physicians 2000).

The setting of physiotherapy has been thoroughly investigated. Physiotherapy can be effectively delivered in either hospital or the community (Forster et al 1999b; Early Supported Discharge Trialists 1999).

From this brief overview, it can be concluded that physiotherapy for people with stroke is effective - the more the better and the sooner the better, but physiotherapy can also be effective long after the initial acute phase of spontaneous recovery, in either hospital or the community as long as it is part of a well-organised, specialist multi-disciplinary service. The next obvious question is to examine the content of physiotherapy to establish which form the most effective interventions. This is a complex issue and is surprisingly difficult to answer. A number of studies have attempted to compare different physiotherapy approaches but no consistent differences have been found (Table 1.1). Although evidence is emerging that approaches that focus on training or facilitating impairments (such as Bobath) may be less effective than a functional or task focused approach (Table 1.1).

Table 1.1 - Trials comparing different physical therapy approaches

(Adapted from Royal College of Physicians 2000)

Authors	Design & Sample	Interventions	Conclusions
Logigan et al 1983	CCT; n=42; acute strokes	Traditional therapy vs Bobath or Rood	No differences found
Dickstein et al 1996	CCT; n=131; acute strokes	Functional or PNF or Bobath	No long-term differences (Bobath slower to regain gait independence)
Lord & Hall 1986	CCT; n=39 in 2 centres	Traditional vs neuromuscular techniques	No differences found
Basmajian et al 1987	RCT; 29 strokes	Bobath vs behavioural approach	No differences found
Jongbloed et al 1989	RCT; n=90 acute strokes (<3/12)	Functional approach vs Bobath/Rood for OT	No differences found
Wagenaar et al 1990	CCT; 7 acute strokes	Bobath vs Brunnstrom	No differences found
Nelson et al 1996	RCT; 26 strokes	Functional tasks vs exercise	Functional task increased supination
Dean & Shepherd 1997	RCT; n=20 strokes (1+yr post-stroke)	Task-specific training in sitting balance vs practice of different tasks	Task-specific training increased performance
Patel et al 1998	CCT; n=184 acute strokes	Impairment-focused approach vs function-focused approach	Similar outcome for disability and placement but LoS 25% shorter in functional unit.
Hesse et al 1995	CCT; 7 strokes >3/12	Bobath vs treadmill training	Treadmill training improved gait; Bobath did not
Richards et al 1993	RCT; n=27 acute strokes	Conventional - early and intensive or less intensive vs early intensive treadmill training	Early intensive treadmill training facilitated gait recovery; no differences the conventional groups
Langhammer & Stranghelle 2000	RCT; n=61 acute strokes	MRP vs Bobath	MRP → shorter LoS, less motor disability, and women had better ADL

RCT = randomised controlled trial; CCT= clinical controlled trial;

PNF = Proprioceptive Neuromuscular facilitation; OT = occupational therapy;

MRP = motor relearning programme; LoS = length of stay

A number of factors have been suggested as contributors to the lack of difference between approaches. Three main factors will be considered here: the intervention; the subjects and the outcome measures.

1.2 The Intervention

An initial explanation for the apparent lack of difference between approaches is that the interventions are essentially directed towards the same goals, and it does not matter what a physiotherapist actually does as long as an intervention is carried out (Ernst 1990). The ultimate aim of all physiotherapy approaches is to maximise motor function, but descriptions of the approaches, and the way in which the approaches achieve improved function appear to vary considerably. These differences include whether the underlying motor problem is perceived to be a disturbance of tone or muscle weakness, whether to use resisted exercise or not, whether to use of guided or independent practice, whether to concentrate on impairment-based `quality of movement ' or functional goals, and how sensory stimulation should be used (Ashburn 1995; Partridge et al 1997).

In Britain, the Bobath approach is predominant (Sackley & Lincoln 1996; Davidson & Waters 2000). The overriding aim of this approach is to inhibit abnormal muscle activity thereby facilitating normal movements, with the assumption that this improves function (Bobath 1990). However, in recent years Bobath therapists have identified ways in which the detail of their current practice differs from that advocated originally by Mrs Bobath (Lennon & Ashburn 2000), and even Bobath experts do not agree on what Bobath now comprises (Mayston 2000a&b; Shelley 2000; Panturin 2001; Langhammer 2001). The way in which Bobath therapy is practised has also been found to vary according to geographic location in the UK and clinical speciality (Lennon & Ashburn 2000; Davidson & Waters 2000). Thus it appears that even the proponents of the Bobath method cannot clearly define and describe the rationale for the method and vary in how they utilise it. Therefore trials comparing approaches may not demonstrate differences, because inter-therapist variation within one intervention group could mask differences between interventions. Studies comparing different approaches have been criticised for a lack of precise information about what was included in the treatment package (Effective Health Bulletin 1992; Ashburn et al 1993; Partridge et al 1997; Kwakkel et al 1998). Without a clear picture of the content of treatment it is difficult identify differences or similarities between approaches.

To clarify the content of physiotherapy practice, several authors have attempted to describe and define the content of therapy for people with stroke by `unpacking the so-called black box'. This has revealed a variety of techniques and activities (Partridge 1994; Ballinger et al 1999; Lennon & Ashburn 2000; Pomeroy et al 2001). However without models to explain the clinical rationale underlying the choice of intervention, it is difficult to understand the descriptions of the interventions. Such models are almost entirely lacking for physiotherapy practice in Britain.

Despite the lack of evidence for complete approaches such as the Bobath approach, there is some evidence for individual interventions. In an extensive review, Pomeroy & Tallis (2000) identified evidence to support the use of the following interventions: progressive exercise therapy, the use equipment, positional feedback, electro-stimulation and EMG feedback. There is however little information about how these interventions should fit into individual treatment programmes. Again, the need to develop a model of stroke physiotherapy is clear - such a model would identify which treatments should be used, how they should be carried out, and provide a rational basis for their use This would provide a clear argument for what physiotherapy methods should be used.

The lack of clear models, or theoretical frameworks, for practice has been identified many times over the last decade as factors limiting the development of effective physiotherapy (Shepherd 1991; Effective Health Bulletin 1992; Partridge 1994; Ashburn 1995; Partridge et al 1997; Wade 1998b). Some suitable methodologies exist and there is a rapidly growing evidence-base available to physiotherapists to define, describe and justify their practice. Thus the time is ripe to develop integrated clinical models of stroke physiotherapy.

1.3 Subjects

Research into physiotherapy for people with stroke, as with other aspects of rehabilitation, is complicated by intra-subject variability. People with stroke vary not only in the severity of the pathology but in the number and severity of impairments. how the impairments translate into disability and eventually handicap, and the rate and extent of recovery (Turner-Stokes 1999). This makes it difficult to define a homogenous population for stroke therapy trials. Intervention studies that do not stratify the subject population are however less likely to achieve a significant result due to variability within the groups, or to involve such small numbers that possible effects are lost in Type II errors or are not generalisable (Effective Health Bulletin 1992; Matyas & Ottenbacher 1993; Turner-Stokes 1999). Many authors have advocated the use of stratification to identify and characterise relatively homogenous groups of people with stroke (Jongbloed 1986; Gresham 1986& 1990; Gladman et al 1992; Kalra 1993). Although various stratification criteria have been suggested none have been widely adopted. The Oxford Community Stroke Project (OCSP) system classifies types of stroke based on the anatomical location of the cerebro-vascular lesion. It is reliable, valid and predictive of recovery (Bamford et al 1991, Lindley et al 1993, Smith & Baer 1999). Disadvantages are that it requires a full neurological examination and although the impairment and disability profiles of people with the same types of lesion are varied (Smith & Baer 1999), two of the syndromes (lacunar and partial anterior circulation infarcts) do not distinguish clear cut recovery patterns (Smith et al 2001). Reding & Potes (1988) advocated a similar but simpler system which is reliable, valid and predictive of outcome (Sanchez-Blanco et al 1999), although it does not include the severity of the symptoms.

Personal clinical experience suggests it is not just the number of impairments that affect the choice of treatment technique but also the severity of impairments. For instance, severe weakness of the leg would limit standing balance, while a milder weakness may limit walking but not standing balance. Different treatment techniques would be used in each instance to treat the weakness. Therefore any stratification needs to discriminate groups which are appropriate for the treatment to be tested. In the present study, people with different degrees of weakness need to be classified so that a homogenous group for whom the treatment is appropriate can be identified.

1.4 Outcome Measures

A final factor identified, as limiting stroke physiotherapy research is the outcome measures used. The main criticisms are in matching appropriate outcome measure with the aims of physiotherapy, sensitivity to detect changes due to physiotherapy: and lack of information about the psychometric properties (Effective Health Bulletin 1992: Ashburn et al 1993; Partridge 1994; Ashburn 1995; Partridge et al 1997; Kwakkel 1998; Turner -Stokes 1999). If outcome measures are to detect change, they need to reflect the aim of the intervention. For example, as the main aim of physiotherapy is to improve motor function, then the outcome measure should be of motor function and not of general ADL. If specific motor impairments are treated then the outcome measures must reflect these and not general disability. Again, this highlights the need for models of clinical practice, as if the most appropriate outcome measure is to be chosen, the aims and objectives of physiotherapy interventions need to be explicit.

The second factor is sensitivity to change. Even if an OM is specific to the physiotherapy activity it may be too insensitive to detect the relevant changes, especially over the short-term. Some ordinal scales require large changes in ability to register a change in score. For example, to increase from a score of one to two on the Rivermead Mobility Index (Collen et al 1991) requires the subject to progress from rolling independently to getting from lying to sitting independently. Such an OM would not detect smaller changes in between these two levels. Other types of OM, such as ratio measures are more sensitive, but may be bedevilled by increased variability or error in performance (Evans et al 1997). They also tend to only be suitable for a narrow range of motor abilities. The timed walk test, for example is only suitable once people can walk safely and independently, which is considered an end-point for rehabilitation in many cases (Tyson & Turner 1999 & 2000).

The final factor concerns psychometric properties. Any OM used to assess the effectiveness of physiotherapy for people with stroke needs to be:

- Reflective of clinical practice
- Suitable for the whole range of stroke severity treated by physiotherapists

- Suitable for use in all the environments in which physiotherapists work
- Sensitive to short-term change
- Reliable
- Valid

The minimal of use equipment and ease of use were identified as requirements for clinical OMs (Sackley & Lincoln 1996; Daley et al 1999). Consequently, these factors would also need to be addressed before an OM could be recommended for clinical use. This is discussed in more detail in Chapter 2.

1.5 The aims of this research

The present study aims to address some of the issues limiting the development of evidence for physiotherapy interventions by considering posture and balance in people with stroke. These aspects of physiotherapy have been chosen because the re-education of 'normal' posture and balance is a cornerstone of physiotherapy for people with stroke in Britain (Bobath 1990; Partridge 1994: Lennon 1996; Ballinger et al 1999; Lennon & Ashburn 2000)

In Chapter 2 a review of published methods of assessing posture and balance to establish ways of measuring posture and balance for use in intervention studies and the clinical setting is undertaken. No methods of measuring posture and balance, which met all of the criteria listed above, were identified. Therefore the first stage in addressing the limitations to development of interventions for posture and balance was to develop a new outcome measure that did meet all of the criteria. As this would need to be drawn from, and reflective of, clinical practice, the experiential literature was reviewed and clinical physiotherapists' views on how to assess posture and balance were sought (Chapter 3). Information from this led to the development of a clinical model for the assessment of posture and balance in people with stroke, which in turn informed the content of a new outcome measure - the Brunel Balance Assessment. This new outcome measure was developed and the psychometric properties tested in Chapter 4. Finally, in Chapter 5 the Brunel Balance Assessment was used to explore the importance of balance following stroke, and to test the original clinical belief that treating balance deficits is essential for recovery. as balance is an essential component of other motor functions.

Chapter 2

Methods of measuring balance and posture

Section 2.1 Introduction

2.1.1. Aim

This chapter reviews the research literature on methods for measuring posture and balance systematically and critically. to identify methods suitable to assess the effectiveness of physiotherapy for posture and balance in people with stroke. The measurement tools need to fulfil a number of criteria (referred to as utility criteria) based on Wade (1992):

- Reliability
- Validity
- Sensitivity to short-term change
- Suitability for a wide range of motor impairments following stroke
- Ease of application
- Suitability for different environmental or social settings
- Portability

Reliability refers to the reproducibility and consistency of score (Bowling 1997). There can be many source of variability - inaccuracy of equipment or test, inconsistency of the tester(s) or subject's performance, or variability between testers. Inter-tester reliability refers to stability in score when several testers perform the same test (Wade 1992). This is important in large studies where there is more than one tester and is relevant to the clinical situation, as any outcome measure needs to provide a similar score whoever performs it. Intra-tester or within-session reliability refers to stability in score, when the same rater and subject repeats the test several times in the same session (Rothstein 1993).Test-retest reliability refers to stability in score when the test is repeated over a period during which change in performance would not normally be expected. (Bowling 1997).

Validity refers to the purpose of the measure and the extent to which it fulfils that purpose (Rothstein 1993). Several aspects need to be considered. Face validity refers to apparent suitability of the measurement for the stated purpose (Dyer

1995). Construct validity refers to the conceptual or theoretical argument that supports the use of a measurement for a specific purpose (Dyer 1995).

Consider a goniometer used as a measure of joint range. This is based on the construct that the angle between two body segments is a representation of the angle at the joint. Content validity considers the content of the measure against the intended aim or purpose – whether it assesses all aspects of the purpose and whether it only measures the content of the construct (Dyer 1995). For instance, an assessment of functional mobility would need to include measures of walking, standing balance, transfer skills, and stairs, but not dressing skills. Assessment of face, construct and content validity is based on theoretical and pragmatic, logical arguments so there are no direct tests. Criterion- related validity refers to how the measure compares to another measure that is already accepted as valid (Bowling 1997). Several aspects of criterion-related validity have been defined. Concurrent validity is often used interchangeably within criterion related validity (Streiner & Norman 1995), but it is also is used to refer to comparison with an established measure of the same construct. Criterion-related validity is used for comparison with an established measure of a different but related construct (Bowling 1997). For instance comparing a balance measure with a measure of activities of daily living. Predictive validity refers to whether the measure can predict outcome (Bowling 1997). For instance whether assessment on admission can predict independence at discharge. Discriminant or prescriptive validity refers to whether a measure can distinguish between different clinical groupings (Rothstein 1993), such as faller or non-fallers (for example, Berg et al 1992a).

Sensitivity to change refers whether the outcome measure detects change in the subjects' performance. The term sensitivity is used interchangeably with the terms 'responsiveness' and 'measurement error'. Measurement error indicates how large a change in score is required to represent a 'true' change in performance. Two components have been identified. Systematic error or bias refers to a general trend for measurements to be different between repeated tests, due for example to the effects of fatigue or practice. Random error arises from inherent biological or mechanical variation or inconsistencies in the measurement protocol, which are not accounted for by standardisation (Atkinson & Neville 1998). This is particularly important with sensitive measures such as interval or ratio measures. Another aspect of sensitivity to change is whether changes at the extremes of the measurement range are identified by the measure. A ceiling effect refers to a limit, after which it is impossible to detect an improvement in performance, while a floor effect is a limit below which deterioration cannot be detected (Streiner & Norman 1995).

Suitability for a wide range of severity

As physiotherapists treat people with a wide range of balance and postural dysfunction, any outcome measure must cover the range of severity that a physiotherapist would be treating.

Ease of Use

The practical aspects of using the outcome measure in a clinical setting need to be considered. Factors to be considered include the time and effort needed to complete the assessment (for the subject and tester), simplicity of use (particularly if any equipment is needed), ease with which the findings can be communicated to other people and capital and running costs.

Suitability for All Settings

Any measure tool needs to be suitable for all the environmental settings in which a physiotherapist works. Physiotherapy is moving increasingly into the community and intermediate care setting (DoH 2000), so any measure of physiotherapy needs to be useable in the home or at the hospital bedside as well as in a treatment gym.

2.1.2 Definition of posture and balance

Before reviewing the literature definitions of posture and balance need to be established. No clear definitions of posture and balance have been generally accepted in the literature. Several terms are used interchangeably: Balance may be used synonymously with equilibrium, postural control and co-ordination activities, while posture is interchanged with postural tone and alignment, while postural sets and key points are ways of describing posture and movement. Although exact descriptions vary, definitions of balance feature the ability to maintain an upright position within the limits of stability or base of support (Carr & Shepherd 1998; Bronstein, Brandt & Woolacott 1996; Shumway-Cook & Woolacott 1995). Posture relates to the alignment or orientation of body segments while maintaining an upright position (Bronstein, Brandt & Woolacott 1996; Shumway-Cook & Woolacott 1995), in which the alignment of the trunk and pelvis plays an important role (Shumway-Cook & Woolacott 1995).

The following definitions are used in the present study:

Balance: the ability to maintain an upright position within the limits of stability or base of support (Carr & Shepherd 1998; Bronstein, Brandt & Woolacott 1996; Shumway-Cook & Woolacott 1995).

Posture: the alignment or orientation of body segments while maintaining an upright posture (Bronstein, Brandt & Woolacott 1996; Shumway-Cook & Woolacott 1995).

2.1.3 Method

Search strategy

CINAHL, Embase and Medline databases were searched using combinations of the keywords listed below. The Cochrane Collection, Core Biomedical Collection and Best Evidence were also used, but no additional relevant literature found.

- Hemiplegia or stroke or cerebro-vascular accident or cerebro-vascular disorder.
- Assessment or measurement or testing
- Balance or equilibrium or postural control
- Sitting
- Standing
- Trunk or pelvis
- Posture

- Walking or mobility or ambulation or gait
- Motor function or control

In addition, references cited in the articles identified from the search, reference lists and bibliographies of related journal articles and books were searched manually.

Limits

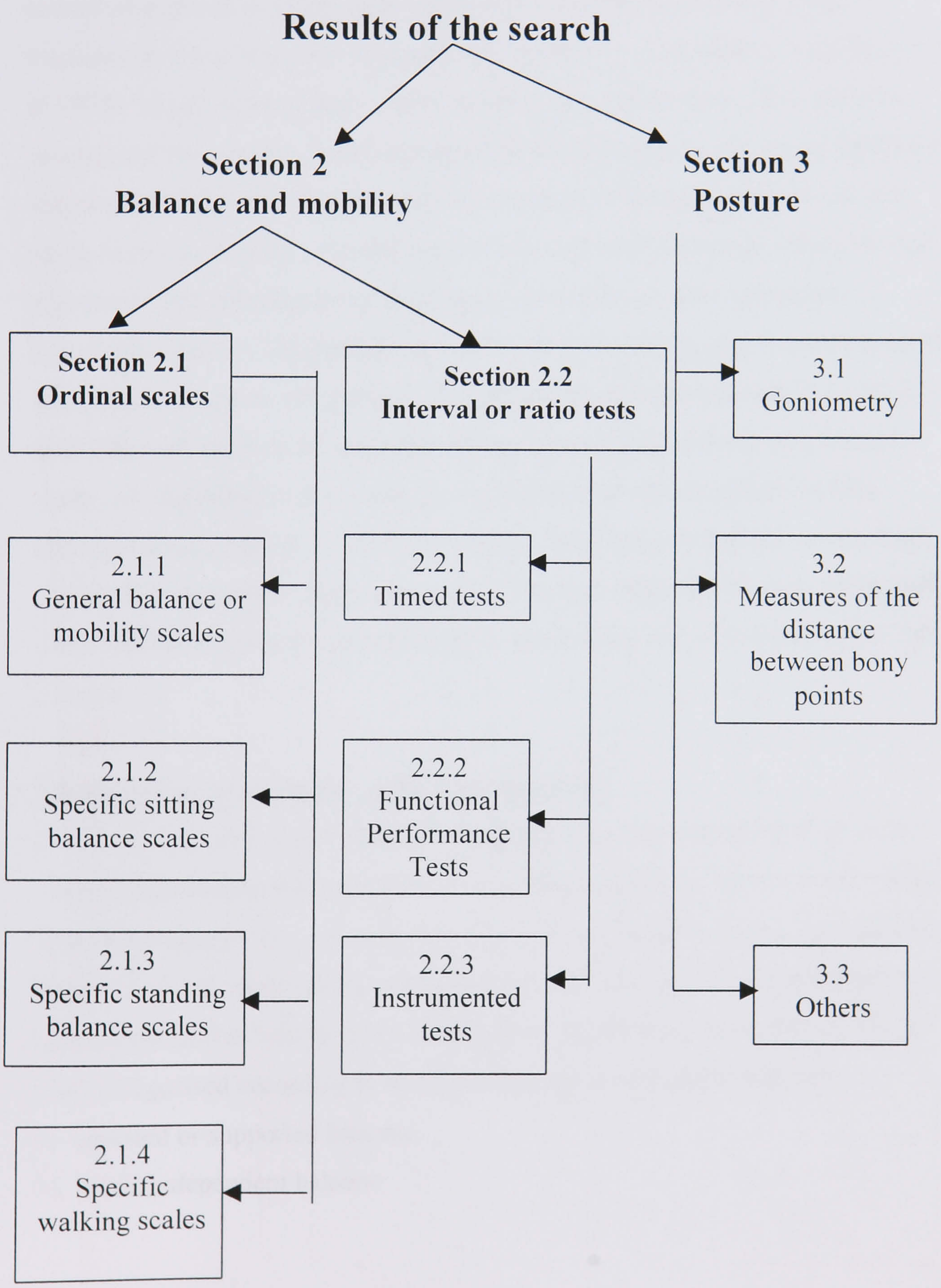
The following limits were used in the searches: English, adult humans and dating from 1960 – 2000. In addition, literature about certain topics were excluded as they would not fulfil the utility criteria, or did not refer to postural or balance dysfunction. They included:

- Tests involving radiography or axial rotation (such as measurement of the Cobb angle used in the assessment of scoliosis), or sophisticated movement analysis systems (Such as VICON or CODA) that would not be suitable for use in all settings.
- Assessments of selective movement in the limbs as a measure of motor function in stroke patients (such as the Motricity Index, Demeurisse et al 1980)
- Assessment of community based disability (such as the mobility section of the Rivermead Extended ADL scale, Gladman et al 1993b).
- Methods of detailed gait analysis involving kinematic, kinetic or temporal-distance parameters (Evans et al 1997 for example)
- Methods of measuring general mobility in the elderly, which have not been developed for use with people with stroke such as the Hierarchical Assessment of Balance and Mobility (Rockwood & McKnight 1995) or the Elderly Mobility Scale (Smith 1994).

Categorisation of results

The search produced a multitude of different assessment methods. The papers were categorised to ease interpretation and comparison. They were initially grouped into assessments of posture or balance, which were then divided according to the type of assessment into ordinal scales or interval/ratio measures. These sections were re-categorised and sub-divided. This categorisation is outlined in Table 2.1 below and the findings of the search are detailed in the sections which follow.

Table 2.1 Flow chart to describe the categorisation of measures to assess posture and balance



PAGE MISSING IN ORIGINAL

as an ordinal scale there was variability in the design of scale, the aspects of balance tested and the tasks performed making comparisons difficult.

Scale Design

Three scale designs were identified. An `index-type` where the subjects' performance on a list of items (presumed to form a hierarchy) was tested on a pass/fail basis, such as the Rivermead Mobility Index (Collen et al 1990). The second design was a `rating-scale' type where a rating scale was applied to a hierarchical list of tasks, for example the Fugl-Meyer Assessment (Fugl-Meyer et al 1975). The number of levels in the rating scales varied (from three points to seven), and the aspects of performance tested also varied. In some assessments the subjects' abilities were rated, in others the degree of independence or whether equipment was required. Another variation was to rate the way in which the task was performed; whether it was performed `normally' or with `abnormal' movement patterns. The third design used a series of sub-scales to break down the tasks and test them in more detail, for example the Motor Assessment Scale (Carr et al 1985) which has sub-scales for sitting, sit-to-stand, walking etc. The sub-scales use standardised tests, such as the ability to maintain a position for a specified time, to assess whether the subject could achieve the task or not. Each sub-scale is intended to form a hierarchy. The type of design of each of the ordinal scales identified from the search is summarised at the end of this Section in Tables 2.3 and 2.4.

Which balance or mobility tasks were assessed?

There was variation in the tasks on which subjects were tested. Mobility tasks ranged from turning the head and rolling to getting up from the floor and walking up and downstairs. The mobility tasks, including balance tasks are summarised in Table 2.2. Even when sitting or standing balance was specifically assessed, there was a great deal of variation in which aspects of balance were assessed. These were categorised according to the difficulty of the tests performed into:

a) assisted or supported balance
b) static independent balance

c) dynamic or self-generated movements

d) responses to external perturbations

(Further details in Tables 2.3.and 2.4).

Although the tests were easily categorised, criteria to pass a task were varied and inconsistent. In particular, the level of performance needed to pass a task was variable and there was no clear rationale for the different choices.

a) Assisted or supported balance. Different types of assistance or support were used with different scales. Most assessments of sitting balance tested whether assistance of another person or support by leaning on the sound arm was required. While for assessments of standing balance a more frequent definition was whether hand support on furniture was necessary.

b) Static independent balance generally indicated an ability to maintain a position without moving and without upper limb support on aids or furniture, or assistance from another person. The duration that the position had to be maintained varied from 10 seconds to 2 minutes for both sitting and standing balance. A number of sitting balance tests had more detailed additions, such as the feet not being in contact with the floor or defining an acceptable posture (Horgan & Finn 1997). A few standing balance tests defined the position of the feet, while others included extra tasks to be performed such as closing the eyes or maintaining single-leg stance (Tinetti et al 1986, Fugl-Meyer et al 1975).

c) Dynamic balance tests assess the ability to maintain a position while performing self-generated movements. Three different types of movement were identified:

- The ability to maintain a position while moving another body segment, i.e. turning the head, looking behind or raising an arm.
- Moving around the base of support, i.e. reaching forwards, sideways or to the floor.
- Changing the base of support (standing balance only), i.e. stepping forwards and backwards, stool stepping, turning 360° or picking something up from the floor.

d) Response to external perturbations was tested most frequently in sitting. This consisted of the ability to withstand external displacement forces pushing the body forwards, backwards or sideways. The application of forces was not described, other than a `sternal nudge', or a `strong or weak force'.

Table 2.2 Mobility tasks assessed in the ordinal scales

	Sitting	Rolling	Lie-to-sit	Standing	Sit-to-stand	Transfers	Walking	Stairs	Other
Berg	✓			✓	✓	✓			Pick up, Turn around360 °
FMA	✓			✓					
BLMA	✓			✓					
PASS	✓	✓	✓	✓	✓				Pick up
MAS (exc UL)	✓	✓	✓		✓		✓		Pick up
RMA-GFS & RMI	✓		✓	✓	✓	✓	✓	✓	Pick up, Bath, Run, Hop
MCA	✓	✓		✓	✓	✓	✓	✓	Bridging, Kneeling, On/off floor
H&F	✓	✓	✓	✓	✓	✓	✓		Lean fwd, Step fwd, back
Ch-McSA		✓	✓	✓		✓	✓		long sitting, on/off floor
Recovery curves	✓	✓	✓		✓	✓	✓		Turn head, stepping
POAM	✓			✓	✓		✓		Turn around
Mobility scale	✓		✓		✓		✓		Bridging
STREAM		✓	✓	✓	✓		✓		Bridging, going downstairs Stepping fwd, backwards, sideways
TCT	✓	✓	✓						

Pick up = picking up an object form the floor, bath = getting in and out of the bath, on/off floor = getting on and off the floor
Long sitting = sitting with both legs out straight

Key to the scales used in this sudy

- Berg = Berg Balance Test (Berg et al 1989)
- FMA = Balance section of the Fugl-Meyer Assessment (Fugl-Meyer et al 1975)
- BLMA = BL Motor Assessment (Lindmark & Hamrin 1988a & b).
- PASS = Postural Assessment Scale for Stroke (Benaim et al 1999)
- MAS (sitting) = sitting sub-scale of the Motor Assessment Scale (Carr et al 1985)
- RMA-GFS = Gross function section of the Rivermead Motor Assessment (Lincoln & Leadbitter 1979)
- RMI = Rivermead Motor Index (Collen et al 1990)
- MCA = Function section of the Motor Club Assessment (Ashburn 1982)
- TCT = Trunk Control Test (Collin & Wade 1990)
- Horgan & Finn = Horgan & Finn Asssessment (Horgan & Finn 1997)
- POAM = Balance section of the Performance Orientated Assessment of Mobility (Tinetti 1986)
- Nieuwboer = Nieuwboer et al (1995)
- S&S = Sandin & Smith (1990)
- Bohannon `86 = Bohannon et al (1986)
- Bohannon `92 = Bohannon (1992)
- Feigin = Feigin et al (1996)
- Ch-McSA = disability section of the Chedoke-McMaster Stroke Assessment (Gowland et al 1993)
- Gabell & Simons = Gabell & Simons (1982)
- Bohannon `93 = Bohannon, Walsh & Joseph 1993
- Recovery curves = Partridge, Edwards & Johnston (1987)
- Mobility Scale = Simondson et al 1996
- STREAM = Stroke Rehabilitation Assessment of Movement (Daley et al 1999)

Table 2.3: Assessment of sitting balance within ordinal scales

Scale	Assisted	Static / Independent	Dynamic/ Self-generated	External Perturbations	Type of scale, No of Items and Points in scale
Berg		Timed. 10s->2min, no support			S, 5, 2
FMA		No support		Yes (unspecified)	R, 3, 3
BLMA	With support	Timed, 10s or 5min		Yes (push, unspecified)	S, 7, 4
PASS	With support	Timed, 10s or 5min			S, 7, 4
MAS (sitting)	With 1 person	10sec, no support, upright posture	Look behind Reach forwards & sideways		S, 6, 2
RMA-GFS		No support, no feet support			I, 1, 2
RMI		10sec, no support			I, 1, 2
MCA	With support or 1 person	60 sec, no support	Reach floor		R, 2, 4
TCT		30sec, no feet support			R, 1, 3
Horgan & Finn		Alignment	Lean forward to floor		R, 2, 4
POAM		Steady, no leaning or sliding			S, 1, 2
Mobility scale		3 mins			S, 5, 6
Nieuwboer		60 sec, Alignment	Lean forwards & sideways		R, 12, 2 or 3 pt
S & S		15sec, No support		Anterior, posterior, lateral nudge - with/out assistance	I, 4, 2
Bohannon 1993	30s With support	30s, No support		Displacement - Strong or weak	I, 1, 5
Feigin	With 1 person	60s	Clasped hands outstretched Lean forward	Resist active displacement forward and lateral	I, 1, 6

With support = using upper limb support (leaning on the sound arm or holding on to furniture), with 1 person = with assistance of another person, I= index type scale, R= rating type scale, S= sub-scale type.
See Key in Table 2.2 for names of scales

Table 2.4: Assessment of standing balance using ordinal scales

	Assisted	Static	Dynamic/ Self-generated	External Perturbations	Other	Type of scale No of Items and Points in scale
Berg	With 1 person Minimal or moderate help	2 mins, no support	Reach forwards		Static eyes open/closed BOS –feet tog. Single stance Dynamic Turn head, Stool step, Stepping, Turn 360°, Pick up	S, 10, 5
FMA	Support	No support			Single stance	R, 4, 4
BLMA	Support	No support- 10s, 1min	Arm movements		Single stance weak & sound leg	S, 7, 4
PASS	No. of people needed to assist	No support - 10s, 1min	Arm movements		Single stance weak & sound leg	S, 12, 4
RMI		10s, No support				I, 1, 2
MCA	Walking aid or chair	30s			Single stance	R, 1, 4
Horgan & Finn		30s				R, 1, 2
Ch-McSA		With 1 person				R, 1, 7
POAM	Walking aid or other	Wide BOS		Sternal nudge x 3	Static - eyes closed, Turn around 360°	S, 5, 2or 3
Gabell & Simons	20s with aid	20s No aid	Turn head x 3		Static - Wide/ narrow/ long BOS Dynamic Turn around 360°	I, 2, 2
Bohannon 1993		30s, Feet apart/ tog			Single stance 60s	I, 5, 2

support = using upper limb support (leaning on the sound arm or holding on to furniture), with 1 person = with assistance of another person, BoS = base of support, tog = together, Pick up = picking up an object form the floor, sternal nudge = applying a displacement force to the sternum, I= index type scale, R= rating type scale, S= sub-scale type.

See Key in Table 2.2 for names of scales

2.2.2 The ordinal scales

2.2.1.1 Scales assessing general motor function

Fourteen scales measuring general motor function including aspects of balance and/or mobility were identified. They are described in detail and compared with their psychometric properties and utility criteria outlined, in Section 2.1

The Berg Balance Scale (Berg) was designed to measure balance in the elderly in clinical or research settings. It has also been tested for use with elderly stroke patients (Table 2.5). Fourteen different activities were tested with a separate 5-point scale for each activity, and the scores for each activity were totalled. It was designed to be used in its entirety, i.e. the activities were not to be used separately. It takes 15-20 minutes to complete. This test is one of the most rigorously developed tests available. It has been found to be very reliable for elderly in-patients and community dwellers, and for both acute and chronic elderly stroke populations (see Table 2.5 for details). Validity has also been extensively tested and content validity was established with a range of health care professionals (Berg et al 1992b). The test has been shown to be a valid predictor of risks of falls and use of walking aids in the elderly. It discriminates between different levels of functional independence and motor recovery and has concurrent validity with other ordinal and instrumented ratio measures of balance and mobility (Table 2.5). It is a more sensitive to change in balance than the Barthel Index (BI) and Fugl-Meyer Assessment (FMA) (Wood-Dauphinee et al 1996). The scaling has not been formally assessed, except internal consistency, which is high. The guidelines state that sit-to-stand and standing balance should be assessed before sitting balance so it is unlikely that the test does form a true hierarchy. Predictive validity has been demonstrated (Wee et al 1999), but sensitivity to change and the measurement error have not been assessed. It may not be sensitive for severely impaired patients as the test of sitting balance only involves static sitting balance.

Table 2.5. Psychometric properties of the Berg Balance Scale

Table A Reliability

	Subjects	Inter-rater	Test-retest
Berg et al 1989	38 elderly with impaired balance	ICC=0.98	ICC=0.99
Berg et al 1995	Acute strokes	n=35, ICC-0.98	n=6, ICC=0.97

Table B Validity

	Subjects	Concurrent	Criterion related	Predictive	Discriminate
Berg et al 1992b	31 elderly	Sway r=0.38-0.55 Timed Up & Go r=0.48 POAM r=0.76			
Berg et al 1989	38 elderly with impaired balance				
Thorbahn & Newton 1996	66 healthy elderly			Risk of falls (with caution)	
Stevenson & Garland 1996	24 chronic strokes	Sway r=0.76-0.81			
Berg et al 1992a	93 rest home residents 60 acute strokes		Barthel r>0.8, FMA r=0.62-0.94	Risk of falls	Use of walking aids, Type of accommodation
Nilsson et al 1998	28 acute strokes	FMA (balance) r=0.77	Cadence r=-0.69 Exertion r=-0.73 FAC r=-63, FMA (total) r=0.62		
Wee et al 1999	128 acute strokes			Length of stay r=-0.6, destination on disch. OR=1.09	

Table C Scale development

		Hierarchy	Measurement Error	Internal consistency
Berg et al 1989	38 elderly with impaired balance			Cronbach's =0.96

The Fugl-Meyer Assessment (FMA, Fugl-Meyer et al (1975) was developed to assess recovery of motor function in stroke patients (Table 2.6). It is based on the methods of Brunnstrom (1966), which are widely used in the USA. This method describes recovery from stroke in a number of sequential stages involving the development and then recovery of 'primitive reflexes and synergies'. There are two main parts to the assessment, one testing sensori-motor impairment and the other a balance scale assessing sitting, standing and single stance. Each position is tested on a 3-point scale. Psychometric testing of the complete test has shown it reliable and valid, but time consuming (30-60mins). There has been little reporting of the predictive or discriminant validity. The scaling properties, hierarchy and internal consistency, have not been reported. Studies assessing the balance section alone (Table 2.6) have found it to be reliable, but the scaling and validity of the Balance section is questioned (Poole & Whitney 1988; Malouin et al 1994; Beckerman et al 1996). On inspection, this may be because automatic reactions to external perturbations in sitting is placed before static standing balance, which does not reflect the relative difficulty of the tasks. The relevance of the patients' ability to resist external perturbations to functional activities is not apparent and has not been formally tested. The inclusion of these tests may cause some lack of validity.

The FMA may also be insensitive for severely impaired people as the easiest item is 'sitting without support' and the next is ability to withstand external perturbations. The FMA is one of the few tests in which measurement error has been assessed, and a change of 4 points on the balance scale (of a total of 14) is required to register a 'true change' (Beckerman et al 1996).

Two further assessments have been developed from the FMA. The Modified chart for Motor Capacity, or **BL Motor Assessment** (Lindmark & Hamarin 1988a&b) has changed the balance section (as well as some of the other sections). This now has a 4-point rather than a 3-point scale, depending on the amount of support required or the length of time the position can be maintained. Seven activities ranging from sitting unsupported to standing on the weak and sound leg are used.

Protective reactions to external perturbations in sitting are assessed subjectively according to their rate of application but there are no details about how the tests should be performed. High internal consistency (Alpha co-efficient = 0.9) for the balance section is reported. As a consequence, the assessment is claimed to be reliable, but it is not further tested. Not surprisingly, there is high concurrent validity with the total FMA score (r=0.95-0.98), and also with the Activity Index (r=0.91-0.92) and the Katz ADL Index (r=0.84-0.87). Greater sensitivity than the FMA is claimed but not demonstrated. Overall, there is no clear advantage in using this assessment rather than the original FMA.

The **Postural Assessment Scale for Stroke** (PASS, Benaim et al 1999) has been developed from the balance section of the FMA. Despite its name, it assesses ability to maintain balance under various conditions and the amount of assistance required. It contains twelve, four-level items assessing ability to maintain and change position in or between lying, sitting, and standing. The levels are based either on the amount of support required to maintain or change the position, or the length of time the position can be maintained. Criterion validity has been found with the FIM (r=0.73), the lower limb Motricity Index (r=0.78), postural sway (r=0.48) and sensation of the lower limb (r=0.45), but not spasticity (r=-0.14). The score at one month successfully predicted transfer and locomotion function at three months and a good degree of internal consistency was found (Cronbach's alpha coefficient =0.95). There was very good inter- and intra- tester reliability for the total score (k=0.99, 0.98 respectively) but it was less consistent for individual items (k=0.45-1). There was a ceiling effect at 3 months. Although a hierarchy was claimed it was not demonstrated and seems unlikely as tasks such as standing on the paretic leg unsupported is tested before sit to stand and stand to sit. Furthermore, items that are probably equally difficult such as standing up and sitting down, sitting up from lying down and lying down from sitting up are tested separately. These would cause plateaux in the hierarchy.

Table 2.6. Psychometric properties of the Fugl-Meyer Assessment

Table A Reliability	**Subjects**	**Inter-rater**	**Test-retest**
*Duncan et al 1983	19 chronic stroke patients		r=0.89-0.98
DiFabio & Badke 1990	5 acute stroke patients	k=0.95	
*Sandford et al 1993	12 acute stroke patients	ICC=0.93	
*Beckerman et al 1996	91 mobile chronic stroke patients		ICC=0.34

Table B Validity	**Subjects**	**Concurrent**	**Criterion-related**	**Discriminate**
Fugl-Meyer et al 1975	28 acute & chronic strokes			
*Dettman et al 1987	15 mobile chronic strokes		Walking speed r=0.61; Sway r=0.52-0.59; Weight distribution r=0.59-0.74	
* Poole & Whitney 1988	30 acute & chronic strokes	Sitting MAS, r=0.28 Walking MAS r=0.64		
*DiFabio & Badke 1990	10 acute stroke patients		SOT r=0.77	
Wood-Dauphinee et al 1990	172 acute strokes	Barthel Index r=0.75-0.85		Total FMA better discriminator of early recovery than the total MAS
* Malouin et al 1994	32 acute strokes	Sitting MAS r=-0.1	Touch r=0.12 Position sense r=-0.1	
* Nilsson et al 1998	28 acute strokes		Berg r=0.77	

Table C Scale development	**Subjects**	**Hierarchy**	**Measurement Error**	**Internal consistency**
*Beckerman et al 1996	91 mobile chronic strokes		4 points	

*Score for the balance section only reported. SOT = Sensory Organisation Test See Key in Table 2.2 for names of scales.

The **Motor Assessment Scale** (Carr et al 1985. Table 2.7) is designed to measure motor recovery of stroke patients. It comprises eight sub-scales of motor function (rolling, lie-to-sit, sitting, sit-to-stand, walking and three measures of upper limb function) and one of general muscle tone. Standing balance is not included. Later modifications discarded the scale measuring tone, as reliability was limited. Each sub-scale, which can be used alone or in combination, is tested on a 6-point scale and the whole assessment takes approximately 30 minutes to complete. Reliability has been established for young and elderly stroke patients undergoing rehabilitation. Content validity is based on the authors' personal experience and choice. The assessment has been validated against other measures of motor performance and functional independence and can detect changes during rehabilitation. The sitting balance scale has been used to predict independence in walking at discharge. The measurement error and scaling (the hierarchy and internal consistency) have not been addressed.

The Rivermead Motor Assessment Scale (RMA) and subsequent **Rivermead Mobility Index (RMI)** assess motor recovery and mobility and are summarised in Table 2.8. The RMA (Lincoln & Leadbitter 1979) consists of three sections, which assess gross function, upper limb abilities, and lower limb and trunk abilities. Only the Gross Function Section (GFS) is considered here. It forms a 13 item hierarchical scale of functional activities ranging from sitting unsupported to hopping on the weak leg. Each item is rated as pass/fail. It has established reliability and scaling, and a 3-point measurement error. It takes about 45 minutes to complete, and has been found reliable and valid when administered verbally which takes 5 minutes (Sackley & Lincoln 1990). It has been validated against other tests of sitting balance and motor impairment and can identify changes during rehabilitation. The RMI was developed from the RMA-GFS as a measure of mobility disability (Collen et al 1990). It is a 15-item hierarchical scale ranging from sitting up over the edge of the bed to running and includes static standing balance. Each item is assessed verbally and scored as pass or fail and it takes approximately 3 minutes to complete. Inter-rater (but not test-retest) reliability and scaling have been established. There is a 2-point measurement error and it

Table 2.7. Psychometric properties of the Motor Assessment Scale

Table A Reliability

	Subjects	Inter-rater	Test-retest
Carr et al 1985	14 chronic strokes	% agreement sitting = 99; walking = 86	Total MAS r=0.98, (0.87-1)
Poole & Whitney 1988	30 acute & chronic strokes	Sitting r=0.99; Walking r=0.99	
Loewen & Anderson 1988	7 acute strokes	Sitting, k=0.56; Walking, k=0.84	Kendall's Tau Sitting, k=0.94-1; Walking k=0.7-1

Table B Validity

	Subjects	Concurrent	Criterion related	Predictive
Poole & Whitney 1988	30 acute & chronic strokes	FMA (balance) Sitting r=0.28, Walking r=0.64		
Loewen & Anderson 1990	50 acute strokes			Sitting balance on admission predicts motor score (r=.83), ADL (r=.82), walking (r=.8) at discharge
Malouin et al 1994	32 acute strokes	FMA (balance): Sitting r=-0.1, Walking r=0.79	Sitting section Touch r=0.38, proprioception, r=0.6	
Nitz & Cage 1995	40 acute strokes			Sitting balance on admission predicts mobility at discharge

Table C Scale development

	Subjects	Measurement Error
Dean & McKay 1992	70 acute strokes	Detects change between admission and discharge
Malouin et al 1994	32 acute strokes	MAS > sensitive than FMA
Nitz & Cage 1995	40 acute strokes	Detects change between admission and discharge

detects change during rehabilitation. It has been validated against other measures of mobility, standing balance and functional independence and for young and elderly, and acute and chronic stroke patients. It can discriminate between fast and slow walkers, patients with and without sensory impairment and patients who did or did not use walking aids (Rossier & Wade 2001). It has also been used to predict walking disability after discharge (Paolucci et al 2001).

A modification of the RMI has recently been published (Lennon & Johnson 2000). This reduces the number of items from fifteen to eight, based on the items physiotherapists chose (at a consensus conference) as essential to evaluate physiotherapy intervention. The scoring was changed from pass/fail for each item to a six-point scale (0-5) depending on the level of assistance the patient required to perform the task. This was in an attempt to improve sensitivity, but subsequent reports showed it did not detect differences between patients better than the original version (Rossier & Wade 2001). Although the other psychometric properties were satisfactory, the modified version offers no clear advantage over the original format.

The **Motor Club Assessment (MCA)** was designed to measure physical deficits and disabilities after stroke and to be suitable for clinical or research use (Ashburn 1982). It was devised from reviews of the literature by a multi-disciplinary group of doctors and physiotherapists interested in stroke research in the late 1970's. The scale is based on a neuro-developmental model of recovery, which is no longer used by British physiotherapists (Lennon & Ashburn 2000). There are two sections; an assessment of motor impairment of the upper and lower limbs and a separate 18-item section testing balance and movement activities. These range from rolling to going up and down stairs. Each item is rated on a 4-point scale. Walking and stairs are timed tests. There has been little testing of the psychometric properties although inter-tester error of one or two points was found in six out of eighteen items in the balance section (Ashburn 1982). The order in which items are tested mean that it is unlikely to form a hierarchy. For instance, getting up and down from the floor is placed before walking. Another criticism is

Table 2.8. Psychometric properties of the Rivermead Motor Assessment and the Rivermead Mobility Index

Table A Reliability

	Subjects	Inter-rater	Test-retest
Lincoln & Leadbitter 1979 (RMA-GFS)	7 acute strokes, 7 testers	ANOVA, No significant differences	N=10, R=0.66
Collen et al 1990 (RMA-GFS)	25 chronic strokes		Kendall's concordance W=0.8
Collen et al 1991 (RMI)	23 neuro pts, inc. 9 strokes	r=0.94; Kendall's tau=0.79	
Lennon & Johmson 2000 (Modified RMI)	30 acute strokes	ICC=0.98	Paired t-test=.73, p<.47

Table B Validity

	Subjects	Content	Concurrent	Criterion related
Sackley & Lincoln 1990 (RMA-GFS verbal)	49 chronic strokes			RMA-GFS (performed) r=0.98
Collen et al 1991 (RMI)	23 neurology patients including 9 strokes	✓	ADL r=0.91; FAC r=0.89; gait speed r=0.82; Standing balance r=0.67; 6min walk r=0.63	
Hsieh C et al 2000 (RMI)	38 acute strokes			BI r=0.67-0.73; Berg r=0.81-0.89
Lennon & Johmson 2000 (Modified RMI)	30 acute strokes	✓		

Table C Scale development

	Subjects	Hierarchy	Measurement Error	Internal consistency
Lincoln & Leadbitter 1979 (RMA-GFS)	51 acute strokes	CR=0.96-0.98; CS=0.91-0.98		
Adams et al 1997a (RMA-GFS)	51 acute strokes	CS=0.81, CR=0.97		
Adams et al 1997b (RMA-GFS)	327 chronic strokes	CR=0.91-0.93; CS=0.59-0.65		
Collen et al 1991 (RMI)	23 neurology patients including 9 strokes	CR=0.93, CS=0.79	Accurate for changes >2 pts	Cronbach's 0.93
Hsieh C et al 2000 (RMI)	38 acute strokes	CR=0.95, CS=0.74-0.79		
Lennon & Johmson 2000 (Modified RMI)	30 acute strokes		Responsiveness; effect size=1.15; error = 4.5 pts	

ADL = Activities of Daily Living, FAC = Functional Ambulation Categories, CR = coefficient of reproducibility, CS = coefficient of scalability the use of the neuro-developmental model, which includes some positions, such as kneeling and half kneeling that, are rarely used in current clinical practice.

Horgan & Finn's (1997) unnamed scale aimed to qualitatively assess recovery of functional mobility and to detect weekly change during rehabilitation. The choice of items and therefore the content validity is not explained but it appears to have face validity. There are 13 functional items ranging from sitting balance to walking and 12 items relating to upper limb function. Each item is scored on a 4-point scale except the walking item which uses a 5-point scale based on the amount of assistance required. The sitting balance and sit to stand scales are more detailed than the other items and include subjective judgements about the posture. Inter-tester (92.5% agreement, kappa=0.82) but not test-retest reliability is reported. Criterion related validity and sensitivity have not been assessed but it appears to detect weekly changes in motor function and changes between admission and discharge. It identified three different levels of severity of motor recovery over a 14-week period and monitored different phases of recovery. It has only been used in an acute rehabilitation setting. The results suggest a floor effect for the most severely affected patients and a ceiling effect for the least severely affected.

The **Chedoke-McMaster Stroke Assessment** is a two-part measure to assess physical impairment and disability after stroke (Gowland et al 1993). The impairment inventory assesses and classifies severity of impairment based on the stages of recovery described by Brunnstrom (1966) and the content and development of the assessment have been clearly described. The disability inventory is designed for use in conjunction with the Functional Independence Measure (FIM) and uses the same 7-point ordinal scale to assess independence / assistance required. This is applied to 10 motor function items and a 5-item walking index, which the authors feel the FIM does not cover in sufficient detail. The gross function section ranges from lie-to-sitting to getting up and down from the floor. Static standing balance is included but sitting balance is not. Inter-rater, intra-rater and test-retest reliability is good (ICC=0.97-0.99). Construct and criterion related validity of the disability section have been established by comparison with the balance section of the Fugl-Meyer Assessment (r=0.68-0.88) and FIM (r=0.83-0.9). It was reported to be more responsive to change than the FIM, although predictive validity and measurement error have not been assessed.

Recovery curves have been suggested as a way to monitor recovery of physical disabilities following stroke by assessing the time taken for patients to achieve functional milestones (Partridge et al 1987; Partridge & Edwards 1988, Partridge et al 1993). Thirteen gross body movements and three upper-limb movements were identified and defined by a group of senior neurological physiotherapists. These ranged from turning the head to walking independently indoors. The patient's ability or inability to perform these was assessed weekly. Inter-tester reliability has been reported (100% agreement) but other aspects of reliability were not, nor has criterion validity or measurement error been assessed. This method has been used for several hundred patients of all ages undergoing rehabilitation following stroke and it appears to detect change over an eight-week treatment period (Partridge & Edwards 1988). The hierarchy is apparent from the curve of the profiles.

Tinetti's **Performance Orientated Assessment of Mobility** was designed to assess mobility and balance skills and the likelihood of falls in the hospitalised and institutionalised elderly (Tinetti 1986). It has not been specifically tested for stroke patients. There are two sections – balance and mobility. In the balance section, performance of nine items ranging from static sitting balance to turning 360° are rated on a 2 or 3-point scale which takes 3 minutes to complete. Good inter-rater reliability is reported (85-90% agreement, Tinetti 1986; k=0.4-1, Ciprany-Dacko et al 1997). Content validity was drawn from literature reviews (Tinetti 1986) and construct and criterion related validity has been tested by comparison with the Berg Balance scale (r=0.91), stride length (r=0.62-.68) and single leg stance (r=0.59-.64) (Berg et al 1992b). It is also reported to predict likelihood of falls in the elderly (Tinetti et al 1988; Lewis 1993). Measurement error has not been established.

likelihood of falls in the elderly (Tinetti et al 1988; Lewis 1993). Measurement error has not been established.

The **Mobility Scale for Acute Stroke Patients** (Simondson et al 1996) is designed to assess the physical status of stroke patients in the acute setting. Five tasks ranging from static sitting to walking on a level surface are assessed. They are rated on a six-point scale depending on the amount of assistance required from others. Test-retest, intra-rater and inter-tester reliability is high (92% agreement, kappa >0.75). Other aspects have not been reported.

The **Stroke Rehabilitation Assessment of Movement Measure** (STREAM, Daley et al 1997 and 1999) is designed to evaluate recovery of voluntary movement and basic mobility post-stroke, particularly in the clinical setting. Content validity was developed in collaboration with physiotherapists to produce three scales (upper limb movements, lower limb movements and basic mobility). The limb movements are scored on a three-point scale, while the basic mobility uses a four-point scale based on the movement pattern used and the amount of assistance required. This gives a maximum score of 30 for the mobility scale. All aspects of reliability and internal consistency for all sections are high (0.98-0.99, and 0.96 respectively for the mobility section). Criterion-related validity has not been tested nor the scaling.

Summary of scales to measure general motor control and mobility

Fourteen scales that assess general mobility, which include aspects of balance were identified. The psychometric properties had not been completely tested for any of the scales and none fulfilled all of the utility criteria. A detailed discussion and summary is found in Section 2.2.1.5.

2.2.1.2 Ordinal scales specifically assessing sitting balance

Seven scales specifically testing sitting balance were identified, including the sitting sub-section of the Motor Assessment Scale (MAS), discussed in section 2.2.1.1. Each test is described below and compared with the utility criteria and psychometric properties.

The **Trunk Control Test** (TCT, Collin & Wade 1990, Franchignoni et al 1997) examines performance of four movements – rolling to the weak and to the sound side, sitting up from a lying position and sitting over the edge of the bed without feet on the floor for 30s. A 3-point scale is used to test each item. Although it is a homogenous scale (Cronbach's alpha=0.83-0.86) the hierarchy has not been assessed and it is unlikely to form a hierarchical scale, as static sitting balance is placed after the more complex tasks of rolling and sitting up from lying. There may also be a floor effect as the first test is to roll to the weak side, which is demanding for people with very limited sitting balance. Inter-rater (r=0.76) reliability has been established and construct validity has been tested by comparison with the FIM (0.71-0.79) and RMA (r=0.79). It is sensitive to change in that it can detect changes over six weeks' rehabilitation at an equal level to the RMA, although measurement error is unknown. It has been used at hospital admission to predict eventual functional and mobility levels at discharge. The test takes less than 5 minutes to complete.

Sandin & Smith (1990) used a series of tests of static and dynamic sitting balance, which were later also used by Morgan (1994). The dynamic tests assessed whether the patient could withstand external displacement forces of "approximately 5-10 lbs force applied anteriorally, laterally and posteriorally". No further details were given about how this was performed or standardised. The subjects' performance for each test was scored on a 4-point scale, considering the amount of assistance required. Construct validity is based on comparison with the Barthel Index (0.7-0.93) and Morgan (1994) reports inter-rater reliability (kappa= 0.8) but details of the testing were not provided. The scale has been used to predict functional independence and mobility at discharge, and to differentiate between levels of sitting balance.

In 1995, **Nieuwboer et al** attempted to develop a rating scale using visual assessment of the quality of posture and balance and trunk activity in five different positions. There were two static positions – normal sitting with feet on the floor and sitting with sound leg crossed over the weak leg (a narrow base of support) and three dynamic activities – leaning to the weak and sound side (on to the elbow) and leaning forwards to touch the feet. Assessment of quality of movement was attempted by making subjective assessments of symmetry and alignment of the trunk and pelvis in the frontal and sagittal planes. Performance was also rated on whether assistance was required. The items were chosen from searches of textbooks on physiotherapy for stroke patients, interviews with three experienced physiotherapists and observation of five stroke patients. Inter-rater reliability was found to be limited particularly for the quality items (balance tasks=0.47-0.65, quality =0.2-0.36). Other aspects of reliability and validity were not tested.

Bohannon and co-workers (1986, 1992) have used two scales to measure sitting balance. In 1986, the subjects' ability to maintain a static sitting posture without upper limb support, foot-floor contact or assistance was assessed on a pass/fail basis. No information about the reliability or validity of this test was given. In 1992, the same definition of sitting balance was used in conjunction with a 5-point scale. This used a mixture of time the patient was able to maintain a position, whether upper limb support was used and whether the patient could recover from undefined "weak" or "strong" external displacement forces. Again, no further information about the psychometric properties was given.

Feigin et al (1996) used items said to be drawn from the sitting balance sections of the MAS and Berg scales. The patient sits on the edge of the bed and reaches forwards with hands clasped to flex the trunk to 90° and straighten up again. Performance is rated on a 5-point scale. 0=unable to sit, 1=required upper limb support, 2=can sit for 60s without upper limb support but can move trunk or limbs, 3= can sit and reach clasped hands outstretched. 4= can perform all of above plus lean forwards and return, 5= can perform all of above plus resist active displacement of balance. No information about the validity or reliability was given, although the authors showed a correlation between sitting balance and gait at discharge and at six and twelve months post-stroke indicating some predictive validity.

Summary

Six scales to specifically test sitting balance were identified, plus the sitting MAS. The psychometric properties had not been completely tested for any of the scales and none fulfilled all of the utility criteria. A detailed summary is found in Section 2.2.1.5.

2.2.1.3 Ordinal scales specifically assessing standing balance

Two scales specifically testing standing balance were identified.

Gabel & Simons (1982) assessed static sitting and then static standing balance with feet positioned with a (defined) wide, narrow or long base. The use of an aid was also noted and whether the patient could perform the test with eyes closed. The ability to withstand internally generated movements - turning the head to look over the shoulder, turning around 360° and standing balance immediately after rising from a chair, were also tested. No rationale for the choice of items was given. Inter-rater reliability was claimed but no details were given. There are no details of the validity of the test.

Bohannon et al (1993) used a seven point ordinal scale to assess standing balance (although scaling was not tested) with items scored on a pass/fail basis. The conditions involve tests (for either 30 or 60 seconds) of bilateral stance with the feet apart (with heels a foot's length apart), together or in single stance. No rationale for the content was given and some items would seem of limited functional relevance to people with stroke and to be beyond the capabilities of the target population - for instance, standing for 60 seconds in single stance.

Reliability of these tests is not convincing. Within-test reliability was reported. 13-18% of subjects changed score within the session, and the ICC for the single stance items were 0.51-0.54 for the weak side. Test-retest reliability was better, reporting ICCs Of 0.6-0.75, but 25% of subjects changed scores from one test and another (1 day apart) which may demonstrate a learning effect. Concurrent validity with walking speed (r=0.66) and mobility (gait FIM, r=0.53) ability was also reported.

The scale was later modified to use times of 30 seconds only and to use single leg stance of the weak or sound leg (Bohannon 1995). Test-retest reliability showed 80-97% agreement and a weighted kappa of 0.905. A correlation between balance and mobility (gait FIM) was found (r=0.86). Bohannon & Leary (1995) had a subject group of mixed disabilities including people with stroke and reported high inter-tester reliability (95% agreement, weighted kappa =0.85) and a correlation between balance and FIM scales for gait (r=0.48-0.77) and transfers (r=0.51-0.68), but less for stairs (r=0.0.23-0.53). They also showed that the ordinal balance scale could detect change over the course of rehabilitation.

Summary

Neither specific standing balance scale fulfilled the utility criteria and the psychometric properties were limited, although the Bohannon's scale was more extensively tested.

2.2.1.4 Ordinal scales specifically assessing walking

Eight ordinal scales specifically assessing walking were identified. Four were sub-scales of the general scales reviewed in Section 2.2.1.1. As the psychometric properties of the walking sections could not be separated from the total scale the content only is described here. A variety of ways to assess walking was used. Some scales assessed the distance that could be covered or the amount of assistance required, while others attempted to test the number and severity of impairments as a measure of the quality of the gait pattern. There is little consistency in the choice of items, nor is there an explicit rationale for their choice. Each scale is described in detail below.

Walking Sub-Scales

The **Motor Assessment Scale** (Carr et al 1985) has a specific walking section. The psychomotor properties showing good reliability and validity have been discussed in section 2.2.1.1. The walking sub-scale has five items which are scored on a pass/fail basis. The items range from standing and stepping with one leg to walking 10m without an aid, pick up an object from the floor and returning in 25s. All the items involve walking indoors so there may be a ceiling effect for the least affected people with stroke.

Horgan & Finn's scale (1997) assesses the amount of assistance the subject requires, firstly from other people and then from a walking aid. Basic reliability has been demonstrated but not validity. As the hardest item is walking without an aid there is probably a ceiling effect, so it would only be suitable for people who are just regaining the ability to walk. The paper claims that the scale can detect weekly changes in performance but this has not been demonstrated.

The **Stroke Rehabilitation Assessment of Movement** (Daley et al 1995, 1997) also has a walking section. Five items are scored on a four-point scale depending on the degree of abnormality of the movement patterns and the amount of aid required. The items are stepping one foot on and off a block, taking three steps backwards and sideways, walking 10m indoors and walking down three steps reciprocally. A broad range of abilities is covered. Reliability but not validity has been demonstrated.

Tinetti's Performance Orientated Assessment of Mobility (Tinetti 1986; Tinetti et al 1988; Berg et al 1992; Lewis 1993; Ciprany-Dacko et al 1997) has mobility as well as a balance section. This scale has a different format to the previous walking scales. The patient is asked to walk across a room or down a hallway at a rapid but safe pace, while the tester seeks the presence of several impairments. This test was originally designed to identify risk of falls in the elderly and the choice of items reflect this rather than the specific impairments of hemiplegic gait e.g. gait hesitancy. step continuity. straightness of the path. step length and height, step symmetry. step width and sway of the trunk. Each item is marked on a two or three point scale.

The **Functional Independence Measure** (FIM Granger et al 1986) was developed as standard measure of disability. There are eleven sections. The mobility section refers to transfers, and the locomotion section, which is discussed here, refers to walking (or using a wheelchair) and stairs. Each section is scored on a seven level scale, which assesses the subject's level of dependence. The items all involve walking indoors ranging from "requiring maximum help to reach 50m and can not achieve 15m independently" to "walking 50m without assistive devices". The stairs section ranges from "unable to go up and downstairs" to "going up and down a flight (12-14) of stairs without a handrail or support". Together these two sections would cover a wide range of abilities. The FIM is reliable (Hamilton et al 1994), valid, and has high internal consistency (Dodds et al 1993). It can detect change in dependence during rehabilitation and changes in carer burden (Granger et al 1993; Hamilton & Granger 1994; Linacre et al 1994); and covers a wide range of abilities. Sensitivity to short-term change is unknown.

Specific Walking Scales

The Functional Ambulation Categories involve a simple six point scale which assesses the amount of assistance needed and then the type of ground the subject can negotiate. It covers a wide range of abilities. is reliable (test-retest kappa=0.36 Collen et al 1990; inter-rater kappa =0.72 Holden et al 1984) and valid (correlation with velocity = 0.67. other temporal-distance parameters of gait r=0.55-0.66 Holden et al 1984). It covers a wide range of abilities, but ability to detect change is unknown.

The **Rivermead Visual Gait Assessment** (RVGA. Lord et al 1998) is intended for day-to-day assessment of gait in people with neurological disease and impaired walking. It uses 18 observations of the trunk and lower limbs and two of the arms and then a four-point scale to quantify the degree of abnormality. Face and content validity were confirmed by discussion with clinical physiotherapists. criterion validity was demonstrated by positive correlations with walking time (r=-0.77), stride length (r=-0.61), Berg balance scale (r=-0.79) and the RMI (r=0.68-0.75). It can detect true change of over 10.5 points, and is moderately reliable. Reliability is better if the global scores are used rather than individual items, particularly displacement of trunk and pelvis. However, the assessment has not been tested specifically with people with stroke (the test group were mainly people with MS and a few people with stroke) and all test subjects needed to be able to walk for at least 10 minutes (with rests). These requirements would make it unsuitable for all but the most able stroke patients.

Summary

Eight scales were identified, most of which were sub-scales of more general tests of motor ability. All the scales had some measure of reliability and most demonstrated some criterion-related validity. Comparisons between them are difficult, as there is wide diversity in the content of items. A detailed discussion is found in Section 2.2.1.5.

2.2.1.5 Discussion of ordinal scales

An ordinal scale should be reliable and valid and by definition, detect improvement, form a hierarchy, and scores should not cluster at either end of the scale (McKensie & Charlson 1986, Wade 1992). Twenty-five ordinal scales of balance or walking were identified from the literature search. All had some information about the psychometric properties but none demonstrated all the properties or the utility criteria as documented in Table 2.9.

Reliability

Most scales had some indication of the reliability but a surprising number, which purport to measure change in status, did not have any reports of test-retest reliability. When formal testing has been undertaken the scales have, by and large, been found reliable. The exceptions were Bohannon et al's (1993) static standing test and Nieuwboer et al's (1995) sitting balance test. Bohannon's test involves demanding tasks such as single stance sustained for 60s (so variability of patient performance could have been the source of error), while Nieuwboer's test relies on visual assessment of quality of movement, which is notoriously unreliable (Wade 1992).

Table 2.9: Summary of the ordinal scales against the utility criteria

	Reliable	Valid as a test of balance	Sensitive to short-term change	Suitable for a wide range of severity	Easy to use/ Suitable for different settings
			General Scales		
Berg	✓	✓	✗	✗	✓
FMA	✓	*	*	✓	✓
MAS	✓	✓	✗	✓	✓
RMA / RMI	✓	✓	*	✓	✓
BLMA	✗	✓	✗	✓	✓
PASS	✓	✓	✗	✓	✓
MCA	✓	✗	✗	✓	✓
HABAM	✓	✓	?	✓	✓
Hogan & Finn (exc. UL)	✓	✗	✗	✓	✓
Ch-McSA	✓	✓	✗	✓	✓
Recovery curves	✓	✗	✗	✓	✓
POAM	✓	✓	✗	*	✓
Mobility Scale	✓	✗	✗	✓	✓
STREAM	✓	✗	✗	✓	✓
			Sitting Scales		
Sitting MAS	✓	✓	✗	✓	✓
TCT	✓	✓	✗	? Floor effect	✓
Nieuwboer 1995	*	✗	✗	✗	✓
Sandin & Smith 1990	✓	✓	✗	✗	✓
Bohannon 1986	✗	✗	✗	✗	✓
Feigin 1996	✗	✓	✗	? Floor effect	✓
Bohannon 1992	✗	✗	✗	✗	✓
			Standing Scales		
Gabel & Simons (1982)	✓	✗	✗	✓	✓
Bohannon et al (1993)	*	✓	✗	✗	✓
Bohannon (1995)	✓	✓	✗	✓	✓
			Walking Scales		
FAC	✓	✓	✗	✓	✓
Tinetti (1986, mobility section)	✓	✓	✗	✗	✓
FIM	✓	✓	✗	✓	✓
RVGA	✓	✓	✗	✗	✓

See Key in Table 2.2 for names of scales. FIM= Functional Independence Measure. RVGA = Rivermead Visual Gait Assessment, ✓ Tested and present, ✗ - Not tested / Unknown * = Tested but not proven

<u>Validity</u>

Most assessments had been tested for at least one type of validity, usually concurrent/criterion related validity and were related to other measures of mobility, balance or motor control and so are generally valid. The only exception being the balance section of the Fugl-Meyer Assessment. The validity of sitting balance tested early after stroke in predicting future recovery has been highlighted and was usually tested rather either concurrent or criterion-related validity

All the scales had face and content validity, although some were based on models not used in Britain, such as the FMA (Fugl-Meyer et al 1975) and Chedoke-McMaster assessment (Gowland et al 1993) which are based on Brunnstrom's stages of recovery and the MCA (Ashburn 1985), which used a neuro-developmental sequence. The external perturbations are used to assess automatic balance reactions in several assessments developed in the USA and Canada. Again, this is rarely used in British clinical practice, where self-generated perturbations or functional tests are more popular. The external perturbation tests also suffered from a lack of detail about how the external forces should be applied and standardised, which would affect reliability.

Another factor limiting the usefulness of some scales was the lack of apparent clinical or functional relevance. For example Bohannon's (1995) standing balance scale tested whether the subject could maintain single stance for 30s, and Gabell & Simons (1982) Balance Coding tested whether standing balance could be maintained with the eyes closed. Again the clinical or functional relevance of testing responses to external perturbations is unclear.

Sensitivity to change and measurement error

Sensitivity to change has also received little attention. When tested, the scales detect change over the long-term such as between admission and discharge from rehabilitation. In most cases, this would reflect a large change in clinical status, and so does not indicate good sensitivity. Horgan & Finn (1997) claimed that their assessment could detect weekly changes in status but this has not been formally assessed. None of the assessments have demonstrated sufficient sensitivity to reflect accurately shorter-term changes during rehabilitation or those due to specific treatment interventions. The relative sensitivity of a few assessments has been addressed by comparison with established assessments – the Berg scale has been found more sensitive than the Barthel Index and FMA (Berg et al 1992a) and the Chedoke-McMaster was more sensitive than the FIM (Gowland et al 1993).

Measurement error (how large a change on the scale is required to reflect a 'true' change in status) has rarely been considered. The tests that have assessed measurement error (the FMA, RMA-GFS and RMI) show a lack of accuracy. The most accurate (the RMI) requires a 2-point change to reflect 'true' change, yet this could indicate the difference between somebody who has standing balance only (a score of 5) and someone who can walk indoors (a score of 7) which is clearly of functional and clinical significance. The rating-scale and sub-scale type assessment might be expected to be more sensitive than the index type scales. The FMA is the only such scale in which measurement error has been assessed. The balance section of the FMA has a measurement error of 4 points (out of a possible score of 14). As each level is rated on a 3-point scale, a change from one level to the next may be purely due to measurement error (Beckermann et al 1996). Methods to improve accuracy and reduce measurement error, such as repeating a test several times and taking an average or best score, have not generally been considered.

Suitability for a wide range of severity - Floor and Ceiling Effects

Another aspect of sensitivity that has received little attention is whether there are floor or ceiling effects i.e. whether scores cluster at either end of the scale. This depends to a large extent on the subject population to which the assessment is applied. For instance, an assessment that only measured sitting balance would show a ceiling effect, if used in the chronic stages post-stroke, as most survivors have good sitting balance in the in the long-term (Partridge et al 1987). The general scales tended to show a floor effect for the most severely impaired people as they did not assess sitting balance and the scales that did test sitting balance tended to restrict this to static sitting balance, which would be insensitive to change. Similarly most of the walking scales were limited to walking indoors, which would show a ceiling effect for the more able people.

Hierarchy / Scaling

The presence of a true hierarchy has two advantages. It means that the whole test does not need to be performed each time the assessment is used and that the final score can give an indication of the subject's ability. Few scales have demonstrated the scaling properties and some clearly do not form a hierarchy in the suggested order of testing. For instance, the Berg scale tests sit-to-stand and standing balance before sitting balance and the TCT tests lie-to-sitting before static sitting. One assessment where the scaling has been tested and found lacking is the balance section of the FMA, where static standing balance is placed after reaction to external perturbations in sitting (Malouin et al 1994). The other main scaling property of an ordinal scale is the internal consistency. Again this had rarely been considered, but where it has been tested the scales were found to be homogenous.

Summary

Ordinal scales have been widely used to assess balance and have generally proved to be reliable and valid. They have added advantages in that they are relatively quick and simple to perform and do not require complex, expensive equipment or specialist training to use. The main disadvantage is lack of sensitivity, particularly for assessing sitting balance. Ratio measures have been used to define different levels and increase standardisation, in an attempt to improve sensitivity and accuracy, but the choice of task and the cut-off point to define different levels lack consistency and functional relevance. Ratio measures also tend to involve some sort of equipment, which can add to time, cost and complexity. The use of a true hierarchical scale is advantageous and can counter these disadvantages, but few of the assessments have a clearly demonstrated hierarchy. The functional or clinical relevance of some tasks in the scales is also questionable.

The Best of the Bunch

None of the assessments comprehensively assess balance and fulfil all of the utility criteria, but the 'best of the bunch' that fulfilled the most of the criteria (see Table 2.9) were:

- The Rivermead Mobility Index as a measure of mobility disability
- The Berg scale as a measure of standing balance
- The Sitting MAS as a measure of sitting balance

The walking scales were too diverse to choose a `best of the bunch'.

Section 2.2.2: Ratio Measures to assess balance and walking

Three types of ratio measure of balance, or walking were identified. These were timed tests, functional performance tests, and tests using instrumented measures. The psychometric properties of each test are considered against the utility criteria.

2.2.2.1 Timed tests

The timed tests were divided into balance and walking tests.

Balance Tests

A number of tests of static standing balance have been developed which assess ability to maintain a static position for a set length of time under varying conditions.

The simplest timed test is the **Romberg test**, which assesses whether a patient can stand with feet together and eyes open and then closed for 60s. The tester observes the amount of body sway and judges whether it increases to an abnormal extent when the eyes are closed (Dornan, et al 1978). A more clinically relevant variation is the **Sharpened Romberg Test** (Briggs et al 1989) which tests the patient's ability to stand with feet in tandem (heel-toe) with arms by their side and eyes open for 60s, it also uses subjective judgement about whether balance is `normal'. Neither of the Romberg tests have reported reliability and given the subjective nature of the test it may be expected to be limited. Another variation of the Romberg test is the **Single Stance Test** (Bohannon et al 1984) which tests the ability to stand on one leg for 30s with eyes open and closed. A further variation notes the length of time the subject can stand on one leg with eyes open and the arms across their chest, which is reported to have inter-rater reliability for the elderly (Gioretti et al 1998). The face validity of these tests is questionable for stroke patients and the reliability is limited (Bohannon et al 1993).

A more recent and more objective development is the **Sensory Organisation Test** (Shumway-Cook & Horak 1986). It is designed to assess ability to maintain static standing balance under a series of conditions that are claimed to test the different sensory factors contributing to balance. The primary aim is to identify the relative contribution of these factors to balance ability, not to assess balance function overall. The original test assesses whether the subject can maintain standing balance for 30s under six different conditions:

- On a firm surface with eyes open (baseline)
- On a firm surface with eyes closed (visual input)
- On a firm surface while wearing a paper dome over the head and face (vestibular input)
- On a soft surface (a piece of medium density foam) with eyes open (reduced proprioceptive input)
- On a soft surface with eyes closed (reduced visual and proprioceptive input)
- On a soft surface while wearing a paper dome over the head and face (reduced vestibular and proprioceptive input)

If the subject is unable to maintain balance for 30s, then the time before moving the feet is taken. Confusingly this test is also known as the Clinical Sensory Integration Test (CSIT) and the 'foam and dome' tests. Another form of the test uses a force platform to assess postural sway under the same conditions and is also called the Sensory Organisation tests (SOT), or the Equitest, after the commercial equipment used to measure it.

Cohen et al (1993) adapted the test for fit elderly people. Testing was for 20s only and they recommended that each condition be repeated three times and mean scores taken. If this is done inter-tester and test-retest reliability is high (k=0.99). Di Fabio and Badke (1990) reported a small pilot study using the SOT with ten stroke patients with standing balance. Good inter-rater reliability was shown (k=

0.77), but there was an apparent floor effect for this group of patients as all the patients could perform the first three conditions for 30 seconds.

Timed Walking Tests

Three timed walking tests were identified. **The 10m-walk test** simply measures the time taken to walk 10m at a comfortable pace with a walking aid if preferred. It is reliable (Holden et al 1984; Wade et al 1987; Collen et al 1990; Rossier & Wade 2001), valid (against the FAC, Holden et al 1984; balance and other measures of gait Bohannon 1987; against the RMI and 2 minute walk test, Rossier & Wade 2001); sensitive (Goldie et al 1996). There is however a measurement error of about 20% largely due to random error of the subjects' performance (Collen et al 1990; Evans et al 1997). It is suitable for different settings: 5m or 3m have been used in community-based studies (Collen et al 1990) and has no floor or ceiling effects until normal speeds are reached.

A variation on the timed walk test is the **Two Minute Walk Test,** when the distance covered in two minutes at the subject's preferred speed is measured (six, ten or twelve minutes have also been used). This tests endurance rather than speed. If the subject is unable to complete the testing period, the distance covered and duration they achieved is then recorded, which should remove any floor effect. The two minute test is reliable, valid (against the 10m walk test and RMI) and can detect differences between people who use aids or not, and who have sensory impairments or not (Rossier & Wade 2001). The six-minute test showed a significant correlation with the RMI (r=0.63, Collen et al 1991)

A final timed walk test is the **Timed Up & Go Test** (Podsiadlo & Richardson 1991), which measures the time taken for a subject to rise from a chair, walk 3m, turn, walk back to the chair and sit back down. It was designed originally as a basic mobility test for the frail elderly and has not been assessed specifically for people with stroke. It is reliable (inter and test-retest reliability kappa=0.99), valid (correlation with the Berg scale r=0.81, walking speed r=-0.61, Barthel Index r=0.78 Podsiadlo & Richardson 1991; Berg et al 1992) and can predict risk of falls (Shumway-Cook et al 2000). Measurement error and sensitivity to change have not been addressed but it appears quick and easy to use and suitable for different settings.

2.2.2.2 Functional Performance Tests

The second category of ratio measures was functional performance tests. These measured aspects of function, such as time, distance or number of repetitions as the subject performed standardised tasks, which all involved self-generated body movements. There were two types; reaching and stepping tests.

Reaching Tests

The reaching tests are based on the **Functional Reach Test** (Duncan et al 1990). This test measures the maximum distance the subject can reach forward beyond arms length in standing, without moving their feet, using a yardstick fixed at shoulder height. The test was originally designed for and extensively tested with the frail elderly. The following have been demonstrated:

- test-retest reliability (ICC=0.8 and 0.88, Duncan et al 1990; Weiner et al 1992) inter-tester reliability (ICC=0.73 Giorgetti et al 1998)
- construct validity (Duncan et al 1990)
- concurrent/criterion-related validity (limits of stability r=0.7, Duncan et al 1990; walk speed r=0.71, ADL r=0.48-0.66, social independence r=0.71Weiner et al 1992)
- predictive validity (risk of falls; Duncan et al 1992)
- sensitivity to change in response to treatment (Weiner et al 1993)
- coefficient of variation of 2.5% (Duncan et al 1990).

Use of the functional reach test has also been tested with people with stroke, but less vigorously. Reliability has not been demonstrated for people with stroke, although they were included in the elderly subjects in Duncan and co-worker's papers (Duncan et al 1990, Weiner et al 1992, Giorgetti et al 1998). For people with stroke, concurrent validity with the step test (r=0.68-0.73, Hill et al 1997), the repeated reach test and the arm raise test and instrumented measures of postural symmetry (r=0.44, 0.43 and 0.66-0.78 respectively) have been reported (Fishmann et al 1997). Berhardt et al (1998) demonstrated responsiveness to change, but there was a moderate ceiling effect for people with stroke undergoing rehabilitation.

A number of variations of the Functional Reach Test have been reported. One study has used **forward reach in sitting** with stroke patients as part of a battery of outcome measures. Change due to treatment over a two-week period was detected, but further information about the reliability and validity was not discussed (Dean & Shepherd 1997). It has also been used with people with spinal cord injury with demonstrated test-retest reliability (0.85-0.94) and discriminate validity and it could discriminate between people with high and low level lesions (Lynch et al 1998).

Brauer et al (1999) used **lateral reach** rather than forwards reach to test the lateral limits of stability in healthy older women. They found a high degree of test-retest reliability (r>0.94) and that concurrent validity with instrumented measures of limits of stability (r=0.33) and movement analysis (0.65). No evidence of use with people with stroke has been found to date.

Two further variations have measured the number of times a subject can reach to a standardised target within a set time limit. Firstly the **repeated reach test** which measures the number of times the subject can reach to a set target within a defined time and secondly the **arm raise test** in which the number of times the subject can lift their sound arm within a set time is counted. These have been tested with stroke patients with high within-session (both r=0.99) and test-retest reliability (r=0.9 and 0.93 respectively) (Goldie et al 1990). Both tests showed systematic error within and between testing sessions although taking the mean value of several tests may overcome this. Fishman et al (1997) reported limited concurrent validity with instrumented measures of postural symmetry for both tests (r=0.02-0.45) and no relationship with postural sway (r=0.00-0.29). There was also a ceiling effect for people with high level of balance function (i.e. walking independently).

Stepping tasks

The **Step Test** was developed to evaluate dynamic single limb stance. It involves stepping one foot on and off a block as quickly and as often as possible within a set time (Hill et al 1996). High test-retest reliability has been found for people with stroke and the healthy elderly (ICC=0.88-0.9), although this can increased further by recording an average or best performance of three trials. It has concurrent validity with functional reach (r=0.68-0.73), gait velocity (r=0.83) and stride length (r=0.82-0.83) and can discriminate between healthy elders and stroke patients (Hill et al 1996). It detects change due to treatment over a four-week period (Hill et al 1997) and does not have a ceiling effect for people with high levels of balance function, although there may be a floor effect for people with limited standing balance (Bernhardt et al 1998)

Goldie et al (1990) used **weight shifting** rather than stepping to assess the ability to transfer weight from one leg to the other in double stance. A variable load monitor was fitted inside the shoe of the subject's weak leg, which was adjusted to make an audible signal when 50% of body weight was transferred on to the weak leg. The number of repetitions within a set time was measured and this was performed in parallel and step stance. High test-retest (r=0.93) and intra-session (r=0.97) reliability has been reported. They recommend that subjects should have one or two practice trials before measurement to reduce performance variability.

2.2.3 Discussion of ratio measures

The ratio measures contained a mixture of tests of balance and walking. The timed balance tests are not encouraging. The psychometric properties have not been widely tested, so that the reliability and validity is unknown in most cases. The timed walk tests are more useful however, especially the 10m walk test which is reliable, valid, sensitive, easy to use, suitable for use in the community and other settings and the tasks performed are clinically and functionally meaningful. The functional performance tests also fulfil these criteria. The equipment used is simple and so they are cheap, need minimal training and, as they are easily portable, they can be used in a variety of settings.

The main disadvantage is the narrow range of subjects for whom each test is suitable – floor and ceiling effects have been demonstrated for different groups of subjects, indicating that each test is only suitable for people who are able to perform the test and find it challenging. For instance people with good functional balance (i.e. were able to walk independently) showed a ceiling effect for the Arm Raise Test but not the Step Test (Hill 1996; Fishman 1997). The Arm Raise Tests assesses static balance in parallel stance as it requires the subject to maintain the position within the base of support, but the Step Test measures dynamic balance in single stance, which is more likely to be impaired in ambulant subjects. The arm raise test may be a more suitable test for people who could not walk, for whom maintaining static standing balance would be challenging.

There are some limitations in the development of these tests at present. Firstly, although test-retest reliability has generally been tested, within-test reliability has not. Where it has been tested, considerable measurement error has been found, although using a best or average score of two or three trials has been suggested to reduce the error (Goldie et al 1990; Hill et al 1996). A related issue is sensitivity to change. Over several weeks changes have been demonstrated, but it may be that performance tests can detect changes over shorter periods, such as individual treatment sessions, but this would be dependent on a clearer knowledge of the measurement error of each test.

Conclusion

Ratio tests, timed walk tests or balance performance tests, fulfil most of the utility criteria. However, no one test is suitable for a broad range of disabilities. It may, however, be possible to develop a battery of tests to cover a wide spectrum of disability.

2.2.3 Instrumented measures to assess balance

The final group of balance measures used instruments or equipment. The aspects of balance measured were postural sway (Section 2.2.3.1), weight distribution (Section 2.2.3.2.) and the ability to withstand external perturbations (Section 2.2.3.3). By far the most common method involved use of platform or force plate tests to measure postural sway.

2.2.3.1 Postural Sway

2.2.3.1.1 Platform Tests

Platform measures of postural sway all used force transducers or strain gauges to assess the ground reaction forces, in anterior-posterior (A/P) and/or lateral directions. These were then used in various ways to calculate the position or movements of the centre of pressure (CoP) within the base of support (BoS). The face validity of these tests is based on the assumption that variability in the centre of pressure (postural sway) is related to functional unsteadiness. Analysis and comparison of the literature were hampered by the huge variability in the studies. The main sources of variability were the tasks subjects performed during testing and the parameters used as a measure of postural sway or steadiness/stability. The different choices are discussed in the following sections and summarised in Tables. Most literature concerns `normal' adults including elderly people but there has been some specific testing of people with stroke.

The Tasks Tested

Many different aspects of performance have been tested which are described below. They are summarised and the references detailed in Table 2.10a & b. There were three main types of task.

1. The ground conditions - The simplest condition was with the platform static. The subject was just asked to maintain a stable position i.e. ordinary standing balance. Subjects were also asked to perform self-generated movements. These comprised two types. In one, subjects were asked to maintain a stable position, while moving another body segment e.g. performing an arm raise. In the other, subjects moved to their limits of support in a specific direction, e.g.

performing a functional reach or shifting weight onto one leg. The final condition was when external perturbations were imposed by moving the platform. In this case the subject was required to maintain position within a moving base of support. The platform could be tilted or shifted linearly forwards, backwards or laterally.

2. The stance conditions - The position of the subjects' feet, or foot (some tests assessed single stance) provided the assessment parameters.
3. Eyes open or closed.

There is little or no rationale or evidence to support the choice of task and none could be clearly identified as preferable. Studies of people with stroke all involved people with a relatively high level of function. They were able to stand for several minutes unaided, sometimes in single stance, or to walk independently. All but one study involved people with chronic stroke. Some of the tests would be of limited importance for a clinically relevant assessment battery - such as eyes closed or reactions to external perturbations. The degree of difficulty also needs to be considered, as it would need to be suitable for people with limited standing balance. Bearing this in mind, the most appropriate tasks would appear to be double or step stance with eyes open, either maintaining a static position or self-generated movements.

<u>The Parameters Assessed</u>

Many different parameters have been used. These are summarised in Table 2.11 and include:

- the average deviation of the position of centre of pressure (CoP) in the x and y axis
- the standard deviation (SD) or the root mean square of the position of CoP in the x and y axis
- the total excursion of the centre of pressure
- the excursion of the centre of pressure in the x and/or y axis
- the total sway based on the area of an ellipse of the x and y axes.

Table 2.10a. The tasks performed in studies of people with stroke

Authors and Equipment	Subjects	Task and ground conditions	Foot position	Eyes
Liston & Brouwer 1996 Balance Master	20 chronic ambulant hemis	Static standing Self-generated movements to LoS	Double	Open, Closed Following visual cue
Fishman et al 1997 Chattx Balance System	20 hemis with independent standing balance <12/12 post-stroke	Static External Perturbations - A/P tilt & shift	Double Step	Open
Dickstein et al 1994 Chattx Balance System	13 ambulant hemis, 6/12 post-stroke 9 age-matched controls	Static; Self generated lateral weight shift to LoS, Lateral external perturbations	Double	Open
Dickstein & Dvir 1993 Chattx Balance System	27 acute hemi with independent standing balance 24 age-matched controls, 13 young adults	Static External perturbations - A/P tilt & shift	Double	Open, Closed
Levine et al 1996 Chattx Balance System	20 non- ambulant hemis	Static, A/P shift & tilt	Double	Open
Di Fabio & Badke 1990	6 ambulant hemis 6 age matched controls	Self-generated movements - weight shift to LoS A/P & lateral	Double	Open
Dettmann et al 1987	15 ambulant chronic male hemis	Static, Self-generated movements - weight shift to LoS, A/P & lateral	Double	Open
Stevenson & Garland 1996	24 chronic hemis with independent standing balance	Static, Self-generated movements- arm raise	Double	Open
Berg et al 1992	31 elderly, including stroke patients	Random external perturbations	Double	Open
Pai et al 1994	14 ambulant hemis >6/12 post-stroke Independent single stance	Self-generated movements - bending knee to raise heel	Single	Open
Winstein et al 1989	61 in-patient hemiplegics with independent standing balance	Static	Double	Open
Ashburn 1996 a & b BPM	23-45 normal subjects, 15 hemiplegics	Static sitting		Open, Closed

LoS = Limits of stability; hemis= hemiplegics; A/P = anterior/ posterior, BPM = Balance Performance Monitor

Table 2.10b. Tasks performed in platform tests using people without stroke

Authors and Equipment	Subjects	Task and ground conditions	Foot position	Eyes
Geurts et al 1993	8 young adults, 8 middle aged adults	Static & Arithmetic task	Double	Open, closed, blurred
Hughes et al 1996	100 community living frail elders	Static	Double, Feet together	Open Closed
Haas & Whitmarsh 1998. BPM	58 young healthy adults	Static	Double	Open
Haas & Burden 2000. BPM	18 young adults	Static	Double, feet together	Open Closed
Hinman 1997 BPM	31 healthy young adults	Static	Double	Open Closed
Le Clair & Riach 1996	25 healthy young adults	Static	Double Heel-toe stance	Open Closed
Goldie et al 1989 Kistler force plate	28 healthy young adults	Static	Double Step standing Heel-toe Single stance All L & R leg	Open Closed
Sackley & Lincoln 1991. BPM	403 healthy adults, 18-87yrs	Static	Double	Open
Duncan et al 1990	156 elderly	Static and forward reaching	Double	Open
Brauer et al 1999	66 healthy elderly women	Self generated lateral reach to LoS	Double	Open

LoS = limits of stability, BPM = Balance Performance Monitor

Table 2.11. Parameters used when testing postural sway

	Parameter Measured and definition
Liston & Brouwer 1996	Postural sway: Area of sway as a % of LoS
Winstein et al 1989	CoP position & stability: Mean position and deviation of CoP in x & y axes
Dettmann et al 1987	CoP position: Mean CoP position in x & y axes
Fishman et al 1997	Postural sway (sway index): Root mean square of deviation of CoP in x & y axes
Di Fabio & Badke 1990	Position of CoP: Excursion of CoP as % of LoS
	Stability of CoP: Density of Position of CoP relative to radii encompassing 30%, 60% and 90% of sway data (bullseye)
Dickstein et al 1994	Postural sway: Standard deviation of position of CoP in x-axis only. Position of CoP: Mean position of COP in x axis only
Dickstein & Dvir 1993	Postural sway: Standard deviation of CoP in x and y axes Position of CoP: Mean position of CoP in x and y axes
Levine et al 1996	Dispersion index: time CoP spends at different distances from mean CoP
Ashburn 1996 a & b	Postural sway: Standard deviation of weight distribution (Sway coefficient)
Duncan et al 1990	Excursion of CoP
Berg et al 1992	Postural sway: root mean square of CoP amplitude and velocity
Pai et al 1994	Postural sway by movement of centre of mass, position of centre of mass relative to BoS
Stevenson & Garland 1996	Postural sway: Speed of excursion of CoP
Geurts et al 1993	Postural sway: root mean square of CoP amplitude and velocity
Hughes et al 1996	Postural sway: Elliptical area from max. movement of CoP in x and y axes
Haas & Whitmarsh 1998	Postural sway: Standard deviation of weight distribution (Sway coefficient)
Haas & Burden 2000	Postural sway: Standard deviation of weight distribution (Sway coefficient)
Hinman 1997	Postural sway: Standard deviation of weight distribution (Sway coefficient)
Le Clair & Riach 1996	Stability: Standard deviation of position of CoP in x and y-axes, average CoP velocity.
	Standard deviation around mean force in x, y & z-axes.
Goldie et al 1989	Stability: CoP in the x and y-axes, forces in x, y and z (vertical) axes.
Sackley & Lincoln 1991	Postural sway: Standard deviation of weight distribution (Sway coefficient)
Brauer et al 1999	Excursion of CoP (sway path): Difference between max. CoP position in y axis and mean CoP position during quiet stance

CoP = Centre of Pressure, LoS = Limits of Stability

Again there is little or no rationale for the choice of parameters. In the main, it appears to depend on what is offered by the equipment used and no parameter emerged as more popular or preferable to another.

Psychomotor properties of tests of postural sway

Despite the extensive use of platform tests to study balance and postural control, there have been relatively few reports of reliability and validity, and the reports that have been published have been mixed (Table 2.12).

Reports of reliability range from very poor to high reliability with no obvious difference in equipment, parameter, technique or subjects to provide explanations. Dickstein & Dvir (1993) found static tasks more consistent than dynamic, but Liston and Brouwer (1996) reported more reliability with dynamic tasks than static. Different levels of impairment in the subjects may explain this contradiction, as Liston & Brouwer (1996) tested chronic ambulant people, while Dickstein & Dvir (1993) tested acute patients with independent standing balance. A number of authors have attributed the poor test-retest reliability to variability of subject performance rather than inaccuracy of the equipment or the rater (Dickstein & Dvir 1993; Pai et al 1994; Liston & Brouwer1996). Taking a mean value of several trials could increase reliability (Haas & Whitmarsh 1998) but this is rarely performed. Measurement error and sensitivity to clinically relevant change have not been addressed but could be crucial if there is a high degree of random error.

Table 2. 12. Psychometric properties of postural sway in studies specifically testing these measures

Authors & Equipment	Reliability	Validity	Subjects and Parameter
Liston & Brouwer 1996 Balance Master	Test-retest. Dynamic ICC =0.84; Static ICC<0.56	Concurrent with Berg scale & gait velocity. Dynamic activities (r=0.48, p<0.5) Static activities not related	Hemis: Self-generated movements
Fishman et al 1997 Chattx Balance System		Concurrent with Functional Reach, arm raise & repeated reach. No correlation found	Hemis: External perturbations
Dickstein & Dvir 1993 Chattx Balance System	Test-retest: Static - moderate r=0.66; Dynamic- poor r=0.09-0.47		Hemis: Static, external perturbations
Dettmann et al 1987		Concurrent with FMA and Barthel score moderate correlations r= 0.52-0.69	Hemis: Static, self-generated movements
Levine et al 1996 Chattx Balance System	Test-retest Weak to moderate. Static ICC=0.75, linear =0.65, tilt =0.8		Hemis: Static, self-generated movements
Stevenson & Garland 1996	Test-retest; No significant difference	Concurrent with Berg scale (r=0.76). BBS > sensitive	Hemis: Static, self-generated movements
Berg et al 1992		Concurrent with Berg r=-0.2-0.67	Elderly, including hemis. Random external perturbations
Pai et al 1994	Sound leg > weak in single stance ICC=0.86: 0.68	Concurrent with velocity (r=0.4), FMA (Leg & Balance) r=0.63-0.81	Hemis: Self-generated movements

Hemis= hemiplegics, BPM= Balance Performance Monitor

Table 2. 12. Psychometric properties of postural sway in studies specifically testing these measures (continued)

Authors & Equipment	Reliability	Validity	Subjects and Parameter
Ashburn 1996 BPM	Test-retest. Hemis < normals		Hemis: Static sitting
Haas & Whitmarsh 1998. BPM	Inter and intra-rater Poor (ICC 0.3-0.5)		Young adults. Static standing
Haas & Burden 2000		Concurrent against Kistler forceplate: Poor validity	Young adults, Static standing
Hinman 1997 BPM	Test-retest (same day retest) ICC =0.62	Concurrent w. movement analysis system. No relationship found (r=0.07)	Young adults. Static standing
Goldie et al 1989 Kistler force plate	Test-retest. Force (0.71-0.85) > reliable than CoP (0.11-0.3)	Concurrent between force and CoP measures. (r=0.16-0.45)	Young adults. Static standing
Geurts et al 1993	Intra-subject reliability - High RMS of A/P CoP velocity most stable		Adults. Static standing
Hughes et al 1996		Concurrent with tests of motor impairments and disability. Weak correlation with strength, none with functional tests or disability	Frail elders, Static standing
Duncan et al 1990	Test-retest poor ICC=0.52. FR > reliable	Concurrent with Functional Reach. (r=0.7). but FR > precise	Elderly: Self-generated movements
Brauer et al 1999	Test-retest high ICC=0.995	Concurrent with lateral reach. (r=0.33)	Self-generated movements

BPM = Balance Performance Monitor. CoP = centre of pressure, hemis= hemiplegics

As with reliability, validity of platform tests has received little attention (Table 2.12). There have been some validity studies, mainly with tests of function (such as Functional Reach, or the Berg Balance Scale). These have failed to show a clear relationship between postural sway and functional balance and mobility. Even concurrent validity between different types of platform is poor, possibly due to the different methods of measuring the changes in centre of pressure (Hinman 1997; Haas & Burden 2000). Postural sway has, however, shown a greater correlation with measures of impairments such as the Fugl-Meyer Assessment and muscle function (Dettmann et al 1987; Pai et al 1994; Hughes et al 1996). It is possible that the platform tests of postural sway assess and therefore identify, limitations in impairment-level postural control mechanisms rather than those used in functional balance activities. This questions the assumption on which the face and construct validity of platform tests of postural sway is based - that the variability of forces produced on the ground equate to the degree of postural sway, which is a representation of the degree of unsteadiness and/ or stability. Other aspects of validity have not been addressed.

When compared, functional balance tests, such as functional reach and the Berg scale, have been found more reliable, valid and sensitive than the platform tests (Duncan et al 1990; Stevenson & Garland 1996), suggesting that tests based on functional performance provide preferable tests of balance. The platform tests have the added disadvantages in that they are expensive, require training to use and are not easily portable.

Platform Tests in sitting.

The platform tests have been used almost exclusively in standing, but two studies have looked at postural sway in sitting. Ashburn (1996a&b) reported using the Balance Performance Monitor to identify differences in postural sway in sitting between hemiplegic and 'normal' subjects' and changes with age but did not address reliability or validity.

2.2.3.1.2 Other instrumented measures of postural sway.

Three other methods of measuring postural sway have been found. These measure movements of the pelvis, which are assumed to measure the position of the centre of gravity (CoG). They give absolute measures of the movement of the CoG, rather than that inferred through the change in forces on the platform. They are based on the assumption that increased movements of the CoG are equated to unsteadiness.

One study (Dickstein et al 1996) used an ultrasound source fixed to the subjects' waist. The signal was picked up by receivers in each of the three dimensions. Sway was calculated in three different ways; the standard deviation of the mean position in the anterio-posterior and lateral directions, the maximum excursion and the speed of excursion. Test-retest reliability was varied (ICC = 0.57-0.95) and measures of speed were considered most reliable. Concurrent validity with platform measures of postural sway were claimed but not demonstrated. It appears to be portable and inexpensive.

The other method uses angular movement in the sagittal (anterio-posterior) plane as a measure of sway (Wright 1971). The instrument (called an ataxiometer) consists of a thread tied to the subject's waist and attached to a mast mounted on a box on the floor. The box contains a double ratchet mechanism to which the mast is attached. Movement of the waist (CoG) is transmitted to the mast and then the ratchet mechanism and any movement of more 3.5 degrees is registered. Unfortunately there are no reports of the reliability or sensitivity. It was used with the frail elderly and a correlation with impaired vibratory sense was found, but not with proprioception, vision or vestibular deprivation. There was also a limited relationship between postural sway and falls in the previous year (Brocklehurst et al 1982).

Magnetometry was used by Fitzgerald et al (1993) and Elliot & Murray (1998). This uses a magnetic field generated by a magnetic coil in a fixed position and a receiver coil worn around the subject's hips. This detects the separation of the coils in the anterio-posterior and lateral directions. The path length was used as best indicator of body sway. The authors report a similar degree of reliability and variability in the subjects' performance to platform tests, but claim superior sensitivity to small changes in sway and greater relevance because it is a direct measure of movement of the CoG.

More recently, accelerometers have been used to assess static and dynamic balance and postural sway (Kamen et al 1998) in young and older fit adults. The equipment differentiated between young and old, healthy elders and those with frequent falls and reliability were high (ICCs=0.75). Other aspects of validity, and measurement error were not reported. It is simple, cheap and suitable for use in the home or other community environments.

2.3.1.3 Discussion of measures of postural sway

There is a plethora of studies measuring postural sway as a means of assessing balance or postural control. However none have convincingly demonstrated their validity. Studies have tended to involve small numbers of subjects with a narrow range of high-level abilities. Relationships between postural sway and other measures of balance or motor disability are moderate at best. In studies where postural sway and functional tests have been compared, the functional tests emerge as superior measures. Reliability is also generally discouraging because of high variability of the subjects' performance during testing. Other disadvantages are that the equipment is usually expensive, requires training to use and is difficult to transport. Measures of postural sway are a poor measure of balance and no clear method of choice has emerged. Functional performance tests have stronger psychometric properties and fulfil more of the utility criteria. They should be used in preference to instrumented tests of postural sway as indicators of balance disability.

2.2.3.2 Measures of Weight Distribution

The other main instrumented measure is weight distribution. In many cases, the same force plate/ platform equipment has been used as that for postural sway, but less sophisticated equipment such as digital bathroom scales have also been used. These are described in detail below and summarised in Table 2.13a and b.

The most common parameter is the mean position of the centre of pressure (CoP) most often expressed in the x (lateral) axis. Percentage of body weight (% body wt) is also used. Reliability has been most extensively tested and generally has been found to be acceptable, but better for lateral weight distribution than the anterior/posterior component. This may be because most of the dynamic tests involved anterior/ posterior movements, so inevitably these movements show greater movement and therefore variability. Validity has received little attention, but when tested, has been found closely related to motor control of the lower limb, moderately related to functional balance tests and weakly related to walking ability.

Most studies have tested weight distribution in standing but a few have looked at sitting balance. Generally test-retest reliability is good with less subject variability in performance than in standing. Only one study (Nichols et al 1996) has looked at validity and found that leaning forward related better to function than static sitting or leaning to the side, and was better at detecting change over time. Another study (Pollock et al 1998) found static sitting symmetry did not change over time, but good sitting balance was an inclusion criterion, so it may be that the subjects' performance reached a ceiling rather than that the equipment failed to detect any change.

Table 2.13a. Details of studies assessing weight distribution

Authors & Equipment	Subjects	Parameters & definition	Task
Bohannon & Waldron 1991	20 independently mobile hemiplegics	Weight (kg)	Static
Caldwell et al 1986 Digital scales	10 ambulant hemiplegics Young & old women	% body weight; Asymmetry index (sound/weak)	Static
Levine et al 1996. Chattx Balance System	20 hemiplegics with independent standing balance.	Position of CoP A/P & lateral	Ext pert: A/P shift & tilt
Dickstein et al 1994 Chattx Balance System	13 ambulant hemiplegics; 9 controls	A/P Weight distribution. Mean location of CoP in x axis	Lateral ext perts & self-generated mvts
Dickstein & Dvir 1993 Chattx Balance System	27 acute hemiplegics; 24 controls; 13 normals	Mean CoP in x & y axes	Ext perts A/P tilt & shift
Fishman et al 1997 Chattx Balance System	20 hemiplegics with independent standing balance	% body weight in x & y axes, symmetry (undefined)	Ext perts A/P shift & tilt
Winstein et al 1989 Force plate	61 in-patient hemiplegics with independent standing balance	% body weight; mean CoP in x & y axes	Static
Pai et al 1994 Force plate	14 ambulant hemiplegics >6/12 post-stroke	Position of CoM in relation to % of width of foot	Self-generated mvts
Sackley 1990. BPM	90 hemiplegics	Mean proportion of body wt on left leg	Static
Nichols et al 1996 Chattx Balance System	4 hemiplegics; 6 controls	% change in weight distribution	Sitting; Static & Self-generated mvts
Ashburn 1996	15 hemiplegics	Weight distribution coefficient (undefined)	Sitting; Static
Pollock et al 1998 Seated platform	28 acute hemiplegics	Symmetry of wt distribution, Symmetry	Sitting; Static
Dean & Shepherd 1997	20 hemiplegics	% body weight	Sitting; Self-generated
Nichols et al 1995 Chattx Balance System	66 young adults	Position of Centre of balance in x & y axes	Ext perts - A/P shift & tilt
Sackley & Lincoln 1991. BPM	403 normals	Mean proportion of body weight on left leg	Static
Haas & Whitmarsh 1998 BPM	58 young healthy adults	% weight distribution	Static
Wing et al 1992 Swayweigh	203 adults	% body weight	Static with & without arm raise

CoP= Centre of Pressure, CoM= Centre of Mass, Ext perts= external perturbations, mvts= movements, A/P = anterior/ posterior

Italics = authors

Table 2.13b. Psychometric properties of tests of weight distribution

Authors & Equipment	Reliability	Validity
Caldwell et al 1986. Digital scales	Accuracy and repeatability claimed but not shown	Not tested
Sackley 1990. BPM		Concurrent symmetry of weight distribution with RMA & ADL - RMA (r=0.45); ADL (r=0.35)
Sackley & Lincoln 1991. BPM		Not tested
Bohannon & Waldron 1991. Digital scales	Within session, high (ICC=.83-.88) Subject variability high - 16%	
Wing et al 1992 Swayweigh		
Dickstein & Dvir 1993. Chattx Balance System	Test-retest; High for lateral weight distribution (r=0.73-0.89); Moderate for A/P weight distribution (r=0.62-0.79)	
Pai et al 1994 Force plate	Test retest; Moderate- high (ICC=0.68-0.86)	FMA leg (r=0.81); FMA balance (r=0.63); gait speed (r=0.4)
Dickstein et al 1994 Chattx Balance System		
Nichols et al 1995 Chattx Balance System	Test retest - Lateral wt distribution ICC >0.6; A/P weight distribution (ICC=0.41-0.56)	Not tested
Nichols et al 1996 Chattx Balance System	Test -retest. Normals high (ICC=0.86-0.97): Hemiplegics moderate/high (ICC=0.5-0.9) Fwd lean can detect change over 1 week but not static balance, or leaning to weak or sound side	Concurrent with FIM Leaning fwd & to weak side related to dressing, transfers & mobility. No relationship with static sitting and leaning to sound side
Ashburn 1996a&b BPM	Weight distribution > repeatable than sway	Not tested
Dean & Shepherd 1996	Not tested	Not tested
Levine et al 1996 Chattx Balance System	Test-retest Lateral weight distribution high, A/P weight distribution poor	
Fishman et al 1997 Chattx Balance System		Concurrent between symmetry of weight distribution & FR & arm raise (r=0.49-0.66)
Winstein et al 1998. Force plate		Not tested
Pollock et al 1998	Not tested	Unable to detect change over time
Haas & Whitmarsh 1998. BPM	Inter, Intra Reliability high (ICC=0.7-0.9). Measurement Error <4%	Not tested

BPM= Balance Performance Monitor, A/P= anterior/posterior. See key in Table 2.2 for names of scales. *Italics =authors*

Discussion of measures of weight distribution

Measuring weight distribution has generally shown reasonable test-retest reliability. It can be performed with simple, cheap equipment but more sophisticated equipment has also been used. In either case, the tests are quick and easy to perform (for subject and tester). There is some evidence of concurrent validity with measures of function and motor impairment although it may be more closely related motor impairment than function. Weight distribution would therefore appear to be a measure of balance meriting further investigation.

2.2.3.3 Other instrumented measures

Measures of ability to withstand external perturbations

The final group of instrumented measures assessed the subjects' ability to withstand external perturbations applied at waist level, in effect directly to the centre of gravity. Three different types of force have been applied

The Postural Stress Test (PST) was developed by Wolfson et al (1986). It used a pulley-weight system attached to the subject's waist to deliver a sudden reproducible backward force. The forces used were 1.5, 3, and 4.5% of body weight. Subjects were given three trials at each weight and rated according to a) the number of trials in which the subject could keep their balance without assistance of another and b) the balance strategy used. The balance strategy was scored according to a nine-point scale developed by the authors based on their own experience and descriptions of equilibrium reactions. Inter-rater reliability of 0.99 is reported, but there are no reports of test-retest reliability, measurement error, or concurrent validity of the scale or performance on the PST. Evidence of validity is very limited. Chandler et al (1990) found differences in performance between elderly fallers and non-fallers but not between healthy elderly and younger adults. Duncan et al (1990) failed to find any correlation between the balance strategy scores and EMG analysis of the patterns of muscle activity in the healthy elderly, or between the balance strategy used during the PST and platform perturbation tests. One study tested people with stroke (Harburn et al 1995). They adapted the original PST by using weights starting at 0.5% of body weight and increasing in 0.5% increments to 4.5% of body weight. Test-retest reliability was high (ICC=0.83-0.93) when the performance over two trials was averaged, and the test distinguished between healthy elderly and high-functioning people with stroke (all the stroke subjects were independent in all ADL, independently mobile and able to stand for at least three minutes).

An adaptation of the PST is the Maximal Load Test (MLT, Lee et al 1988), which assesses ability to withstand static forces applied in backward, forward and laterally. A pulley-weight system was attached to a waist belt. The initial weight was 1lb and extra weights were added at 1% of body weight at a slow and constant speed until a) assistance was required to prevent a fall, b) the subjects used their hands to maintain their balance or c) their feet were raised from the floor. Performance was compared to timed tests of standing balance in double, tandem and single stance and a `qualitative assessment of tilting reactions'. This was applied to healthy elderly and stroke subjects. No information about reliability is given. There was a moderate correlation between the maximal load applied anteriorly and posteriorly and clinical balance scores (r=0.66, 0.68), but not for laterally applied forces. The maximal force distinguished between hemiplegics and healthy subjects, but not between young and older subject, nor between weak and sound side in people with hemiplegia.

The final variation on the PST tested lateral forces applied to the hip (pushes) using a `computer-assisted electronically driven actuator system' (known as BERTI, Wing et al 1993). It was tested on people with hemiplegia and age-matched control subjects. Kinematic response to the application and removal of the pushing force was studied and compared to a balance impairment rating scale. No details of reliability are provided. No convincing relationship between response to the forces and the balance rating was found. Differences in performance between hemiplegic patients and control subjects were identified, but not between the hemiplegic patients' weak and sound sides.
None of these external perturbation tests have been extensively developed, and none offer promise that they may fulfil the utility criteria for use in people with

stroke. The face validity of external perturbations as a measure of functional balance is not apparent: One rarely comes across destabilising forces (whether pulling or pushing) applied at the hip in everyday life. There are also ethical concerns about the use of destabilising forces with people who are known to be vulnerable to falls. The use of 'spotters' or a safety harness to prevent a fall addresses the issue of safety, but it does not address the affront to the subjects' dignity that such a test could cause. Although not extensively tested, reliability could be poor as subjects' responses to destabilising forces would be variable and initial evidence of validity is not encouraging. In the light of this plus the lack of portability, the use of external destabilising forces will not be pursued further in this study.

Instrumented Measures to Assess Walking

There are a number of instrumented ways to measure walking, from simple footprint analysis to sophisticated 3-D movement and kinetic analysis. The aim of these tests is detailed analysis of the gait pattern rather than to test balance ability. Furthermore, they are not easily portable and therefore unsuitable for use in different environments, so they will not be considered further in this review.

Discussion of instrumented measures

Many instrumented methods to assess balance have been identified but few have been satisfactorily tested and fulfil the utility criteria. Platform measures of postural sway have limited reliability and validity as a measure of balance ability and reactions to external perturbations have been insufficiently tested. Measurement of weight distribution does show more promise however. Here there is reasonable evidence of reliability and validity, the equipment is simple to use although it can be expensive and difficulty to move, but simple 'low-tech' methods have been suggested which merit further investigation.

2.3.2 Conclusions of methods to assess balance

Numerous methods of assessing balance have been reviewed but few have been rigorously developed and none meet all the utility criteria. Several methods do show promise however. Ordinal scales are reliable. valid. easy to use and suitable for a wide range of abilities and different settings, but they lack sensitivity to short-term change. Ratio measures are also reliable and valid, easy to use. suitable for different settings and may be sensitive to short-term change. but each test is only suitable for a narrow range of abilities. Instrumented balance measures were disappointing, but one method that shows promise is to use digital bathroom scales as a simple `low-tech' method to measure weight distribution.

Section 2.3 Methods to measure posture

2.3.1 Introduction

There is a large body of literature about the measurement of posture, most of which relates to orthopaedic and rheumatology interests in back pain. No ordinal scales of posture were identified, although some of the balance scales reviewed in Section 2 included observation of posture as part of the balance assessment. Nieuwboer et al (1995) attempted to produce an ordinal scale including observations of sitting posture but reliability was unsatisfactory. Horgan & Finn (1997) reported good inter-rater reliability for a scale that mixed balance, posture and impairments with disabilities. but they did not specifically address the reliability of the postural parts of the scale. Carr et al (1999) developed a bedside checklist to assist positioning. However the reliability was unsatisfactory particularly for positions of the trunk and head.

The most frequent measures of posture are ratio and these have been divided into three sections

1. Methods using goniometry
2. Methods measuring the distance between bony points or anatomical landmarks
3. Other methods.

A number of methods, which would evidently not meet the utility criteria were rejected. This excluded methods using x-rays, or other sophisticated equipment such as Moire Fringe Analysis or contourography, computerised motion analysis systems such as Coda or Vicon and photographic or video methods requiring digitisation. Literature concerning the measurement of scoliosis was also excluded.

2.3.2 Goniometry

Goniometers measure the angle between a body segment and the horizontal or vertical, or the angle between two body parts. There are a number of different designs, but an inclinometer has been shown to be the type of choice (Burdett et al 1986), so it is reviewed in detail here.

An inclinometer consists of a fluid-filled disc with a bubble incorporated in it. The disc is graduated in 1-degree intervals and can be rotated relative to the base to allow the device to be zeroed. The disc can be mounted on to a small base plate or two-point contact to allow application to the body part. It is positioned on the body part to be measured and then the angle relative to horizontal or vertical (depending on the position required) is read from the position of the bubble. It is small (pocket-sized), light and relatively cheap. A variation on this fluid or bubble inclinometer is the gravity inclinometer, which has a gravity dependent needle rather than a bubble. Digital and computerised versions are also described in the literature although they do not offer any clear advantage over the manual version (Chiarello & Savidge 1993).

The inclinometer has been most extensively used to measure pelvic and lumbar spine posture and movement, mainly in healthy young people or people with low back pain. No studies using inclinometers in people with stroke were found but the technique was reviewed in the hope that techniques, which could be adapted for people with stroke, might be discovered. A number of different postures and movements have been assessed. The most common technique (the dual inclinometer technique) uses two inclinometers to assess movements of the lower back. One inclinometer is positioned over the sacrum (S1), and another over the thoraco-lumbar (T12/11) junction. The patient is then asked to bend forwards as far as possible. The sacral inclinometer measures pelvic (hip) movement, while the thoraco-lumbar inclinometer is said to measure total movement. The range of movement at the lumbar spine is then calculated by subtracting the score for the pelvic/hip movement from the total movement. This technique has been very widely used as a test of range of movement of the back and has been adopted as the preferred measure of impairment due to low back pain by the American Medical Association. Consequently, it is used by worker compensation organisations in the USA, Australia and New Zealand to assess disability and impairment (Nitske et al 1999).

Criterion-related validity has been demonstrated by comparison with x-ray measurement (Mayer et al 1984; Saur et al 1996) and methods measuring the distance between bony points (Reynolds 1975; Burdett et al 1986; Gill et al 1988). It differentiates between people with and without back-pain and detects change during rehabilitation (Mayer et al 1984). Despite the adoption of this technique by the worker compensation organisations its validity as a test of disability is questionable (Nattrass et al 1999).

Reliability (inter-tester, within-session and test-retest) of quiet standing and flexion is reported to be good to excellent (Reynolds et al 1975; Burdett et al 1986; Keeley et al 1986; Gill et al 1988; Rondinelli et al 1992; Chiarello & Savidge 1993; Saur et al 1996). Lumbar spine extension shows less accuracy and reliability due to the small movement involved (Reynolds et al 1975; Keeley et al 1986; Gill et al 1988; Chiarello & Savidge 1993; Saur et al 1996). The technique has also been found satisfactory to measure lateral flexion (Reynolds et al 1975). The inclinometers are accurate, with a measurement error of 5 degrees (Mayer et al 1995 &1997). The main source of error is variability in testing technique. Accuracy and consistency in identifying the bony landmarks is the primary problem, although this improves with training and practice.

Single inclinometers have also been used to measure pelvic and lumbar spine movements in flexion, extension and lateral/ side flexion while standing and in flexion while sitting. This is cheaper and simpler than the dual inclinometer technique, as the tester only has to take one reading at a time. It is highly reliable (Reynolds 1975; Mayer et al 1984; Burdet et al 1986; Mellin 1986; Mellin et al 1991; Newton & Waddell 1991; Rondinelli et al 1992) and concurrent validity with x-rays and 'finger-to-floor test has been reported (Newton & Waddell 1991).

Inclinometers have also been successfully used to measure cervical range of movement. A frame or adjustable straps are used to hold inclinometers in position to measure flexion/extension, side flexion and rotation. This had good inter- and intra- tester reliability and was more reliable and accurate than a universal goniometer or visual estimation (Tucci et al 1986; Klaber-Moffat et al 1989; Youdas et al 1991; Love et al 1998).

Summary

Inclinometers have been widely used to measure lumbar, pelvic and cervical position and movement most frequently in quiet standing and forward flexion to assess low back pain. Inclinometers are reliable and are a valid measure of position, movement and impairment, but not disability in people with low back pain. There are no published reports using an inclinometer to measure posture in people with stroke.

2.3.3 Methods of measuring the distance between bony points.

The distance between bony points has been used as an indication of range of movement. One study considered the distance between bony points in stroke patients. In a pilot study using non-stroke subjects, Taylor et al (1995) used the distance between the inferior angle scapula and the second and eighth thoracic spine and between the anterior superior iliac spine and acromium in sitting as measures of muscle tone. The validity of this assumption was not tested, but these measures could also be said to be a measure of posture per se. The authors developed a novel variation of a body caliper to account for subject obesity and showed good reliability when measuring static resting postures. However some subjects complained of backache during the testing session.

Other studies have been carried out with a view to measuring orthopaedic and rheumatoid conditions particularly of the lumbar spine. They were included in the review as it was hoped they might reveal methods that could be used for people with stroke. The best known is the Modified Shober technique, which measures the distance between markers 10cm above, and 5cm below the posterior superior iliac spines in static standing. The subject then bends forwards (touching toes) as far as possible and the distance between the two points is measured again. The difference between the two readings is considered a measure of the lumbar spine movement. High reliability is reported by Fitzgerald et al (1983), Gill et al (1988) but not by Reynolds (1975). There is no obvious explanation for the contradictory results. In a development of this technique, Moll et al (1972) used skin distraction to assess lateral flexion with good reliability and a plumb-line to assess extension (Moll et al 1972).

Another widely used method is to measure the distance from the finger-tips to floor during forward bending to measure flexion (of the hips and back combined) or during side bending as a measure of side flexion. This is reliable for flexion but less satisfactory to measure extension and rotation (Frost et al 1982; Gill et al 1988; Newton & Waddell 1991).

Gajdosik et al (1985) measured the distance between the anterior or posterior iliac spines and the floor to assess pelvic tilt. Reliability for the resting position and anterior pelvic tilt were good, but was less so for posterior pelvic tilt. Allowing the subject to practise the movements required before testing was found to be important in improving reliability.

Grimmer (1996) assessed neck posture in the sagittal plane by measuring the distance between C7 and a fixed position vertical ruler, between the ear and the same vertical ruler. This was used to assess the degree of cervical excursion (poking chin posture). Test-retest reliability was acceptable and although a number of postural patterns were identified they were not related to incidence of neck pain.

Summary

Measuring the distance between bony points is a reliable way of assessing range of movement, although the validity of these tests has received little attention. Most methods would be unsuitable for use with people with stroke as they involve movements that are neither relevant nor feasible, such as bending forward to touch the toes. If movement is measured (rather than static resting posture), the subjects are required to sustain the position while the tester takes the measurements. This would be difficult for subjects who found the movement challenging and impractical if several measurements were being taken. Like goniometry, accurate identification of the bony landmarks is the main source of error and several authors report difficulties using this method with overweight subjects or people with loose skin, both of which are common in people with stroke.

In conclusion, measuring the distance between bony points is a possibility for assessing static posture and single movements. However, it would be impractical for assessing posture during the dynamic activities and using for multiple measures.

2.3.4 Other methods

The final section included a number of novel methods developed to measure posture of the pelvic or lumbar spine. A flexi-rule or flexi-curve is a pliable rule used to measure the contours, lordoses, or kyphosis of the back, mainly lumbar lordosis. Once moulded to the contours of the spine, the flexi-rule retains the shape while it is traced on to paper and then the angle between pre-set landmarks can be calculated geometrically. A flexi-rule has been used to measure static standing posture, extension in prone and flexion during forward leaning in sitting (nose to knees). It has only been used to measure sagittal movements of the lumbar spine; measurements of side flexion have not been reported. It has good intra-tester reliability (Walker et al 1987; Youdas et al 1995 & 1996) and validity with x-ray measurements (Burton 1986; Hart & Rose 1986). Using a flexi-rule does, however require the subject to maintain their position while the ruler is applied which could be difficult for subjects with poor balance performing activities they found challenging.

Two studies reported the use of a bubble level (a type of spirit level). Gross et al (1998) used a bubble level attached to body calipers to measure whether the pelvis was level in the sagittal and coronal planes. If the pelvis was not level it was not possible to obtain an objective measure of the degree of tilt however. Donahue et al (1996) used a bubble level with a ruler attached to measure lateral shift in people with back pain, but reliability proved unsatisfactory.

Finally, McLean et al (1996) used a projected shadow and a plumb line to measure list of the trunk (the distance between surface markings of T12 and S1) in standing. The plumb line proved to be the method of choice with acceptable reliability and precision. However it is doubtful whether the method could be used in sitting when there is a much greater flexion, as the plumb line may not hang freely.

Summary
These 'other methods' show little promise as a measure of posture in people with stroke, as reliability is limited and validity is relatively untested. The testing protocols are only suitable for static resting postures.

2.3.5 Discussion of methods to assess posture

Despite the importance attached to rehabilitating 'normal' posture by physiotherapists (Bobath 1990), the assessment of posture in people with stroke has received little attention in the literature. Attempts to use ordinal scales to assess posture have proved unreliable and so instrumented ratio measures may be more appropriate.

Of the instrumented measures, inclinometers are the most promising, but they have not been used in people with stroke and several obstacles would need to be overcome. The main source of measurement error has been identified as inconsistency and inaccuracy in identifying bony landmarks. This is more difficult in people who are overweight or have loose skin, so this would need attention. In the literature, the tester holds the inclinometer(s) and is assumed to move it (them) accurately with the subject. This is another potential source of error, especially for stroke patients whose movements can be unpredictable. Another problem is how the measurements could be taken when the tester needs both hands free to assist the subject if necessary. In the literature, the subjects are able to move and sustain positions without help while several measurements are taken. If this were necessary for people with stroke it would exclude people who found a position challenging. These are the people a clinician would most want to measure! Attaching the inclinometer to a belt or harness may overcome the need for the tester to hold the inclinometer and improve accuracy of placement, but a method of taking several different readings quickly, possibly in different planes, would still need to be overcome. A digital or computerised version of the inclinometer would enable a reading to be taken from any angle, or for the reading to be recorded and read later rather than in real time, once the patient is in a safe relaxed position.

Conclusion

The inclinometer is reliable, valid as a measure of posture, accurate and easy to use. If suitable equipment can be found, and suitable protocols developed, it could cover a wide range of abilities and different settings, so its use merits further investigation in people with stroke.

Section 2.4: Conclusions of methods to assess balance and posture

Many methods of assessing balance and posture have been reviewed but few have been rigorously developed and none meet all the utility criteria. However, several methods show promise. It may be possible to combine ordinal scales with a battery of ratio measures to develop a balance assessment that has the advantages of both types of assessment. For assessment of posture, inclinometers may be adaptable for use with people with stroke and simple `low-tech' methods of measuring weight distribution may also be viable.

As no method to assess balance, which meet the utility criteria for balance assessments have been found, it is necessary to develop a new assessment. The lack of apparent clinical and functional relevance and a clear rationale for the content of some assessments has been criticised. If a new assessment is to be developed, it must be clearly drawn from clinical practice. The first step in developing a new outcome measure is to develop such a model. In the next Chapter the process of assessing balance in the clinical setting will be explored, to inform the content of the new assessment.

Chapter 3

How physiotherapists assess balance and posture post-stroke

Introduction

In Chapter 2 a systematic review of research literature failed to identify methods of assessing posture and balance (P&B) that met the utility criteria for use as an outcome measure in the clinical or research setting. Therefore a new outcome measures needs to be developed. If an outcome measure is to assess physiotherapy interventions then it needs to reflect accurately clinical practice. Despite the lack of suitable published outcome measures in the literature, physiotherapists carry out assessments every day. Therefore the aim of this chapter was to examine how physiotherapists assessed posture and balance in the clinical setting, in order to inform the content of a new outcome measure. This was done in two parts - Section 3.1 reviewed the experiential literature on physiotherapy for people with stroke to identify recommended methods of assessing P&B. Then in Section 3.2 focus groups with neurological physiotherapists explored how they assessed P&B in their daily clinical practice.

SECTION 3.1: THE EXPERIENTIAL LITERATURE

3.1.1 Introduction

The experiential literature on physiotherapy for stroke was reviewed to establish the process of assessment and means of measuring P&B. Assessment was defined as the process and means of identifying problems (impairments or disabilities) with the patient's posture and balance.

The experiential literature was defined as books published between 1980 and 1997 by physiotherapists working in stroke rehabilitation who had achieved peer acclaim. Peer acclaim was defined as citation in the Physiotherapy Index (literature search database) and publication in peer reviewed journals, or appearance in the recommended reading lists of the under- or post- graduate neurological rehabilitation modules at Brunel University. The aim was to capture the literature that physiotherapists' (student or qualified) used to inform their practice. Two previous surveys have shown that neurological physiotherapists regularly read professional literature and identified the most commonly read authors and texts (Nilsson & Nordholm 1992; Carr et al 1994). The books identified in these surveys are included in this review. 1980 was chosen as chronological cut-off point so that the literature would be reasonably current.

Thirteen books were identified, they are listed below.

- Bobath B. (1990) Adult Hemiplegia; Evaluation and Treatment (3rd Ed.) Butterworth Heineman, Oxford
- Carr J & Shepherd R (1980) Physiotherapy for Disorders of the Brain. Heinneman Medical, London
- Carr J & Shepherd R (1987) A Motor Relearning Programme for Stroke (2nd Ed) Butterworth Heinemann, Oxford
- Cotton E & Kinsman R (1987) Conductive Education for Adult Hemiplegia. Churchill Livingstone, London
- Davies P. (1985) Steps to Follow. Springer-Verlag London
- Davies P. (1990) Right in the middle. Springer-Verlag London
- Davies P. (1994) Starting Again. Springer-Verlag London
- Edwards S. (1996) Neurological Physiotherapy: A Problem-Solving Approach. Churchill Livingstone, London
- Johnstone M. (1987) Restoration of Motor Function in the Stroke Patient; A Physiotherapist's Approach. Churchill Livingstone, Edinburgh.
- Lynch M & Grisigno V (1991). Stroke and Head Injuries: A Guide for Patients, Families, Friends and Carers. J Murray, London
- Partridge C. (1994). Evaluation of Physiotherapy for People with Stroke. Kings Fund, London
- Sawner K & LeVigne J (1992) Brunnstrom's Movement Therapy in Hemiplegia: A neurophysiological approach. (2nd ed). J. JB Lippincott Co. Philadelphia
- Shumway-Cook A & Woolacott M (1995). Motor Control Theory & Practical Applications. Williams & Wilkins, Baltimore

In these books, physiotherapists have written about their approaches to treating people with stroke. These are based on their clinical experience and they have, to varying extents, developed theoretical models and used the evidence-base to support their clinical methods.

3.1.2 Method

A thematic content analysis was used to analyse the textbooks to establish the process of assessment and means of measuring P&B. Content analysis is a process by which information is systematically sifted and categorised to give meaning and explanation to the data (Mayas & Pope 1995). It is most commonly used with field notes or transcriptions from qualitative methods such as observational studies or interviews. However the general method can be applied to written material as well. Firstly to familiarise herself with the literature the author read all the books. Then the index of each book was used to identify material relevant to assessment and means of measuring P&B. The following hierarchy was used:

1. Chapters or sections on assessment of P&B.
2. Chapters or sections on general assessment
3. Chapters or sections on treatment of posture and/or balance
4. If none of the above were included in the text, a search of the index using the key words - assessment, measurement, posture and balance were used to identify sections or paragraphs.

The books were then re-read to check that no relevant material had been missed. Having identified all the relevant material, this was read to identify themes relating to the aims of the review (how physiotherapists assessed P&B and to identify means of measuring P&B). Four initial themes were identified. Two, the purpose of assessment and the process of assessment related to the first aim – the process of assessment. The other two, recommended methods of measuring P&B and important physical features (postures, movement or actions) related to ways of measuring posture and balance.

The material pertaining to each of the themes was coded (using coloured pens), then each theme was analysed again and further themes identified, coded and categorised. The purpose and process of assessment was not categorised any further but the material regarding methods of measurement was more detailed and was sub-divided. The recommended methods of measurement were categorised into formal outcome measures, informal objective measures and informal subjective measures (section 3.1.3.3). The `important features` section was the largest and most complex of all the categories. This was categorised into positions and/or activities used, position or alignment of body segments, weight distribution and muscle activity (section 3.1.3.4). Finally, the author re-read the material to check whether there were any other themes or issues that had not been identified but nothing further was found.

3.1.3. Results and Discussion

The way in which assessment was tackled varied between the books. Two books Shumway-Cooke and Woolacott (1995) and Carr & Shepherd (1987) had chapters on postural control or balance problems, which included assessment and treatment of these problems (level 1). Carr & Shepherd (1980), Johnstone (1983) and Davies (1985) had general chapters on `assessment of the stroke patient' and Bobath (1990) had a chapter on evaluation of motor patterns. All of these included information about posture and balance assessment. A further four books (Carr & Shepherd 1987; Davies 1990; Partridge 1994; Edwards 1996) did not consider assessment explicitly but they consistently identified certain physical features as important when describing treatment techniques. By implication those features could be important for inclusion in an assessment to identify deficits that require, or respond to treatment, so these texts were analysed to identify those important features. These are summarised with the other texts in Section 3.1.3.4. Three books (Cotton & Kinsman 1987; Lynch & Grisigno 1991; Davies 1994) did not consider assessment at all in any way and these were removed from further analysis. The four themes (identified in section 3.1.2 above) are considered in detail below.

3.1.3.1 The purpose of assessment

The most frequent reason for assessment was to "identify the patient's problems" (Carr & Shepherd 1980; Davies 1985; Johnstone 1987; Sawner & LeVigne 1992; Shumway-Cook & Woolacott 1995). Other reasons were to plan treatment and monitor progress and the effects of therapy (Carr & Shepherd 1980; Davies 1985; Shumway-Cook & Woolacott 1995). An exception was Bobath (1990) who felt that the main purpose was to assess the quality of the patient's motor patterns.

3.1.3.2. The process of assessment

The scope of assessment varied and different aspects of motor function - balance, posture, and functional mobility were generally entwined with selective movement of the limbs and upper limb function into an assessment of the patient's 'overall status'. There was a strong emphasis on assessment of impairment rather than disability. Balance abilities and functional mobility received little attention other than a basis for positions in which to assess posture or selective movement of the limbs. Essentially, most authors suggested observational analysis of movement impairments in a number of different positions. The positions and activities varied but sitting, standing and walking were the most common themes (Table 3.1). The patient's abilities - what s/he could do and his/her level of independence received very little attention in comparison to the movement patterns s/he may or may not exhibit while doing a movement.

The process of assessment was rarely explicit. Shumway-Cook & Woolacott (1995) alone gave a specific structure to assessing P&B. Assessment of postural control was divided into a) the ability to perform functional tasks requiring postural control, b) sensory and motor strategies used to maintain position in space and c) the motor, sensory and cognitive impairments that constrain control. Methods of assessing each area were suggested and four formal outcome measures recommended (section 3.1.3.3).

None of the other texts had a specific model to define and explain the process, or how the findings were used to inform physiotherapy interventions. Carr & Shepherd (1980) presented a general model of physiotherapy for people with stroke which included aspects of assessment, in particular the need to identify missing or abnormal components of a movement by comparison that expected for a 'normal person'. Carr & Shepherd (1987) offered a hierarchy of increasingly demanding balance tasks to be tested (progressing from maintaining static postures to dynamic activities moving around the base of support and then to automatic activities). The emphasis was on analysing (by observation) how components of the movement pattern used to achieve the task differed from normal, but they did not suggest formal outcome measures.

The other books (Davies 1985; Johnston 1987; Bobath 1990; Sawner & LeVigne 1992) offered informal methods of assessment - essentially descriptions of the process and checklists of features to look out for. Some informal objective measures were suggested (summarised in Section 3.1.3.3). The content of the checklists were based on the authors' experience and preference and no rationale or theoretical basis to their content was explained. Tasks, movements, and positions to be tested were suggested but information on why they were included, how they should be judged, or how the results could be used to inform treatment was lacking. All the checklists emphasised careful observation of the patient's movements, based on comparison with that expected of a 'normal' person. The content of those checklists are summarised in Section 3.1.3.4.

The validity of three checklists (Johnstone 1987; Bobath 1990; Sawner & LeVigne 1992) is questionable. Johnstone (1987) and Sawner & LeVigne's (1992) lists are based on a neuro-developmental sequence of balance tasks. This includes activities such as rolling, kneeling and crawling before standing and stepping, which is now thought inaccurate for stroke patients and is not used in Britain (Lennon & Ashburn 2000). In addition, Sawner & LaVigne's list (1992) is based on Brunnstrom's stages of recovery (1966), which encourage the development of spasticity. This is contrary to physiotherapy as practised in Britain (Sackley & Lincoln 1996; Lennon & Ashburn 2000) and so limits the content validity of Sawner & LaVigne's checklist for use in Britain. Bobath's list (1990) is of reactions (mainly in the limbs) to external perturbations in different positions. The use of reactions to external perturbations as an assessment or treatment tool has been criticised, as most balance activity is proactive and self-initiated rather than reactive to an external force. So it is thought that reactions to external perturbations do not reflect the balance reactions for normal every-day activities (Carr & Shepherd 1987; Shumway-Cook & Woolacott 1995). This limits the

Bobath's checklist as an assessment of a patient's functional capabilities. An additional problem is the level of ability required for patients to complete the checklist, s/he needs to be able to sit, kneel, stand and step under a variety of conditions, so it is only suitable for people with a mild or moderate hemiplegia. There are no recommendations about what should be done for people who could not complete the checklist.

3.1.3.3 Recommended methods for measuring posture and balance

There was much more information about how to measure P&B then the process of assessment. Shumway-Cooke & Woolacott (1995) was the only text to use formal outcome measures. The following measurement methods were identified from the texts. The books in which they were recommended in italics.

Formal objective methods to measure P&B

Functional balance measures:

- The functional reach test (Duncan et al 1990): *Shumway-Cooke & Woolacott (1995)*
- Timed Get up and Go test (Podsiadlo & Richardson 1991): *Shumway-Cooke & Woolacott (1995)*
- Berg Balance Scale (Berg 1993): *Shumway-Cooke & Woolacott (1995)*
- Tinetti Balance and Mobility Scale (1986): *Shumway-Cooke & Woolacott (1995)*

Measures of muscle activity:

- Muscle tone (Modified Ashworth scale, Bohannon & Smith 1987) *Shumway-Cooke & Woolacott (1995)*
- Muscle activity/strength: manual muscle testing (Oxford scale, *Wade & Langton-Hewer 1987*), dynamometer: *Shumway-Cooke & Woolacott (1995)*

Informal objective methods

- Symmetry/ Vertical Alignment: Plumb line: *Shumway-Cooke & Woolacott (1995)*
- Weight Distribution: weighing scales, force plates: *Shumway-Cooke & Woolacott (1995)*

- Joint range/ Muscle length: goniometry, distance between bony points,
- resistance to movement or stretch *Shumway-Cooke & Woolacott (1995)*

Informal subjective methods

- Alignment: Observation of position and movement of body segments, comparison with normal: *Carr & Shepherd (1980),Johnstone (1983), Davies (1985), Carr & Shepherd (1987), Davies (1990), Bobath (1990), Shumway-Cooke & Woolacott (1995), Edwards (1996),*
- Muscle tone: observation of muscle, observation, of alignment of body segments, 'feel of movement', resistance to movement or stretch: *Johnstone (1983) Davies (1985). Carr & Shepherd (1987), Bobath (1990)*
- Muscle activity/strength: observation of muscle, observation of alignment of body segments: *Carr & Shepherd (1987) Shumway-Cooke & Woolacott (1995)*
- Joint range/ Muscle length: goniometry, distance between bony points, resistance to movement or stretch: *Johnstone (1983), Davies (1985), Bobath (1990), Shumway-Cooke & Woolacott (1995)*

3.1.3.4 The content of measures of P&B: The important features

Each text was searched for physical features that were consistently highlighted as being important for assessment or treatment of P&B. Four important 'features' were identified:

- The positions and/or activities used
- Position or alignment of body segments
- Weight distribution
- Muscle activity

All the texts dealt with assessment of people with stroke except Edwards (1996), who described the analysis of normal posture and movement in detail as a basis for comparison with stroke (or other neurological) patients. She does not however explain how this should be done, or how the results of any assessment should be interpreted.

<u>The positions or activities used.</u>

As highlighted in Section 3.1.3.2 assessment of balance received little attention in comparison to the emphasis on motor impairments. However, some positions and activities were consistently suggested (Table3.1) sitting and standing were the most common positions.

Table 3.1 Positions and activities used in the assessment of P&B.

	Rolling	Lie<-> sit	Sitting	Standing	STS/ Transfers	Walking	Other
Bobath			✓	✓		✓	Kneeling Single stance Prone lying
Carr & Shepherd (1980; 1987)		✓	✓	✓	✓	✓	
Davies (1985; 1990)	✓	✓	✓	✓	✓	✓	Stairs On/off floor
Edwards		✓	✓	✓	✓	✓	
Johnston	✓		✓	✓			Kneeling
Partridge	✓	✓	✓	✓	✓	✓	
Sawner & LeVigne			✓	✓		✓	
Shumway-Cook & Woolacott			✓	✓			

STS = Sit-to-stand

<u>Position and alignment of body segments</u>

In assessing posture and motor impairments, all authors recommended observation and subjective analysis of alignment and movement of body segments, most frequently of the trunk, pelvis and symmetry (Table 3.2). Carr & Shepherd (1987) and Edwards (1996) gave particular emphasis to comparison with the patient's expected norm. They, and Shumway-Cook & Woolacott (1995) described the essential features of normal posture and movement for comparison, and Carr & Shepherd (1987) described common abnormalities to look for. Bobath (1990), Johnstone (1983) and Sawner & LaVigne (1992) did not measure position and alignment of body segments.

Table 3.2. Alignment of body segments in the assessment of P&B

	Head / Neck	Trunk	Pelvis	Scapula	Symmetry
Bobath					
Carr & Shepherd (1980)	✓	✓	✓		✓
Carr & Shepherd (1987)	✓	✓	✓		✓
Davies 1985	✓	✓	✓	✓	✓
Davies 1990	✓	✓	✓		
Edwards	✓	✓	✓	✓	✓
Johnston					
Partridge	✓	✓	✓	✓	✓
Sawner & LeVigne					
Shumway-Cook & Woolacott	✓	✓	✓	✓	✓

Weight distribution

Weight distribution was also consistently referred to, but less often than alignment. All authors referred to weight distribution, primarily lateral weight distribution through the buttocks (in sitting) or feet (in standing), but the position of the centre of gravity (CoG) within the base of support was also an issue. Carr & Shepherd (1980; 1987) and Shumway-Cooke & Woolacott (1995) considered postural sway although no means of assessing it was suggested. Johnston (1987), Bobath (1990), and Sawner & LaVigne (1992) did not consider weight distribution at all.

Muscle Activity

The assessment of muscle activity was more controversial. Several authors (Carr & Shepherd 1980; Davies 1985; Johnstone 1985; Bobath 1990) reject the assessment of muscle strength, recruitment and joint range as they state that muscle activity is dominated by spasticity which is task and position specific. This, they feel, invalidates the assessment of muscle activity other than tone, but suggestions of how to assess tone were all subjective and vaguely worded. However more recent texts (Carr & Shepherd 1987; Davies 1990; Shumway-Cook & Woolacott 1995; Edwards 1996) have accepted the role and importance of muscle strength, and/or range for function.

Table 3.3 Assessment of muscle activity

	Muscle Tone	Muscle Length	Muscle Activity
Bobath	✓		✓
Carr & Shepherd (1980; 1987)		✓	✓
Davies (1985; 1990)	✓	✓	✓
Edwards	✓	✓	✓
Johnston	✓		✓
Partridge		✓	✓
Sawner &LeVigne			✓
Shumway-Cook & Woolacott	✓	✓	

3.1.4 Conclusions of the review of experiential literature

The experiential literature was reviewed to analyse methods and means of assessing P&B in stroke patients. Overall, the review revealed a lack of structure and coherent rationale to the assessment methods. There was general agreement that the main purpose of assessment was to identify the patient's problems, which was used to inform treatment, but there was little consensus on the process of assessment or means of measurement.

Most measurement was informal and relied on the physiotherapist's subjective observational analysis. By far the most important and consistent aspect of measurement was observation of the alignment and movement of body segments in different positions and activities (sitting and standing being the most common). These were compared with the expected norm for the patient to identify disabilities and impairments. The trunk and pelvis were the body segments most frequently cited. Objective methods to assess this were lacking. Assessment of function received little attention in comparison to motor impairments.

SECTION 3.2: HOW PHYSIOTHERAPISTS ASSESS POSTURE AND BALANCE

Introduction

The review of experiential literature had thrown little light on how physiotherapists assess posture and balance. Coherence about the process of assessment and how to measure it was lacking. A few objective and subjective measures were recommended and 'important features' were identified but none of these could be used as a basis of a new outcome measure. Consequently, practising clinical neurological physiotherapists were consulted about how they assessed posture and balance. Specifically the aim of this study was a) to identify how physiotherapists measured posture and balance during clinical assessment and b) to develop a clinical model about the assessment process.

3.2.1 Method

Study design

A number of qualitative methods were considered to fulfil the aims of this study. Observational designs were rejected because the physiotherapists' views were wanted rather than a researcher's interpretation of their actions. Questionnaires were rejected as specific information was wanted and it was anticipated that some interaction between researcher and participant would be needed to elicit this. One-to-one interviews were rejected as another study reported that physiotherapists in a similar situation tended to give what they thought was the 'right answer' rather than their actual practice with one-to-one interviews (Lennon & Ashburn 2000). So a method involving a group interaction was considered most appropriate. Focus groups are a style of group interview whereby data arises from the interaction and discourse generated by a group discussion (Morgan 1997) and this method was used to obtain a snapshot of current clinical practice.

There were however some ways in which the aim of this study differed from those in which focus groups are usually used. Focus groups are usually used to explore participants' views, feeling and opinions where a group interaction would elicit information which would not be forthcoming using other methods (Bowling 1997). For example in a recent study Lennon & Ashburn (2000) used focus groups to explore physiotherapists' understanding of how the Bobath concept had changed in the last decade and their opinion of the main theoretical assumptions. In the present study the information desired was much more factual: A report of what physiotherapists did when assessing and measuring P&B was the desired outcome rather than their opinion or knowledge of the background to why they did it.

In focus groups, discussion topics are introduced by the researcher who acts as a facilitator. The level of moderation is determined by the desire for control of the sessions and the topic under discussion (Morgan 1997). In this case, the information sought from the focus groups was specific and very factual, so photographs of a 'typical' stroke patient were used to focus the participants to the questions and to 'get the ball rolling' (Appendix I). Firstly a photograph of the patient in a relaxed sitting position was shown and participants were asked: *"What features of posture and balance would you note if you were assessing this patient?"* Once the features for relaxed sitting were exhausted, photographs of the patient taking weight through his sound side, his weak side and then in a standing position were also shown. The participants were again asked what features of posture and balance they would note if assessing the patient. The choice of positions was based on personal clinical experience and the positions and activities used in the experiential texts.

Data collection in focus groups is usually undertaken having an open discussion during which the conversation is taped or notes are written by an uninvolved observer (Sims 1996). These are later transcribed and details of the discourse and contributions of individuals are included in the data collection and subsequent analysis. This method was initially the plan for this study. However, in a pilot group this produced a stilted conversation because it was difficult for participants to remember what factors had been identified as the discussion progressed. Another problem was variability in the terminology used to describe the features that they observed. This made it difficult to know which suggestions were new features and which were the same features described in a different way. To overcome these problems the facilitator acted as scribe and wrote suggested items on a flip chart during the discussion. This produced an instant form of data and gave the participants an indicator of 'where they had got up to'. It also increased transparency of the data (as the participants could see what was written) and allowed instant feedback on accuracy and completeness. Participants reminded the facilitator if any of their suggestions had not been included, or if items were rejected after further discussion. The facilitator could also clarify the terminology used without appearing to challenge their views. This method produced a list of features and other jotted notes at the end of the session and these, rather than transcriptions, were used as the data for subsequent analysis. The data did not include information about the group interaction or contribution of individuals beyond brief informal notes made by the facilitator after each session. However, this was thought a reasonable price to pay to obtain richer data reporting the physiotherapists' practice which was the aim of the study.

Each group lasted about one hour and was arranged at the participants' convenience at their workplace. Initially an outline of the overall project, specific objectives of the focus group and how it would run was given. Participants were reassured that it was their opinion that was sought, that there were no right or wrong answers, and no judgement was attached to their contributions. It was explained that there was a lot of variability in the terminology and jargon used in this area and so the facilitator might ask for clarification about some of the language they used, but this was merely for clarification and to ensure that their views were accurately represented.

Subjects

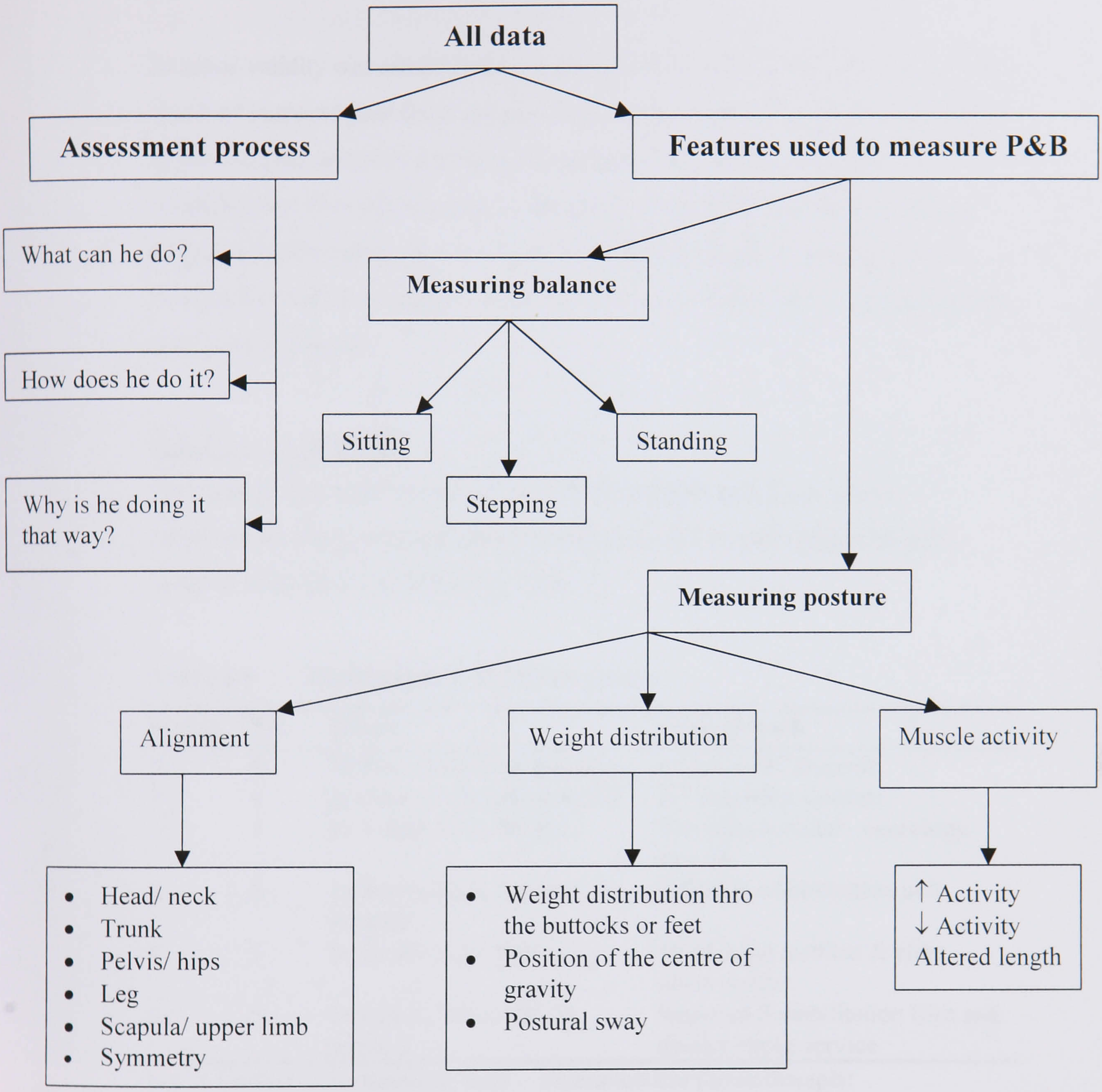
Groups of local physiotherapists with an interest in neurology were invited to take part in the focus groups. Subjects were recruited from physiotherapists on the MSc in Neurological Rehabilitation at Brunel University and neurological physiotherapists working in four different local NHS trusts. Lead clinicians of the local units were contacted by letter to explain the project and to invite them (and their staff) to take part. The MSc students were approached personally. A time to hold the focus group was then arranged at their convenience so attendance was assumed to indicate consent. Participants were working at Senior II level or above, so they could be assumed to bring a reasonable degree of experience and knowledge to the discussion and were working regularly with stroke patients (at least once a week).

Analysis

Content analysis was used to analyse the data from the focus groups. At the end of the focus groups the data consisted of notes, suggestions, words and phrases about assessing and measuring P&B, presented in a haphazard fashion on sheets of flip chart paper. To analyse this, the contents of the flip charts were transferred to a word processing package (Word). Then the data were sorted by grouping similar words and phrases together. This was repeated to sort the material into more and more specific categories. The categorisation that this analysis produced is illustrated in Figure 3.1

The data were combined into one large analysis rather than considering the content of each focus group and each individual's contribution in detail, which is the usual method (Bowling 1997). This was because the aim of the study was to get an overview of physiotherapy practice. The desired outcome was a factual account of what the physiotherapists did during clinical assessment. Their feelings, opinions or interactions between individuals – the usual issues addressed in a focus group (Bowling 1997), were not the concern in this study. The groups were created according to the participants' workplace merely for convenience, rather than grouping similar people together which is the more common practice (Bowling 1997). The groups included people with a mixed level of seniority who worked in a variety of clinical areas (Table 3.4). They were similar in that they all worked with stroke patients at senior II level or above but no more detailed grouping was attempted.

Figure 3.1 Flow chart to illustrate the categorisation of data from the focus groups about how physiotherapists assess posture and balance in people with stroke.



<u>Reliability and validity of the analysis</u>

Reliability of the content analysis was checked by three independent assessors (all lecturers at Brunel University). They went through the data categorisation and noted whether they agreed or disagreed with the categorisation. Percentage agreement for each sub-category was calculated.

To check internal validity, the author returned to two of the original focus groups and presented the results. Participants were asked whether the results reflected the original focus group and their clinical practice.

External validity was established by presenting the results to a group who were not involved in the original focus groups. This was done pragmatically at a conference about outcome measures for physiotherapists working in neurological rehabilitation. The audience was invited and were all senior physiotherapists in specialist rehabilitation units throughout the United Kingdom. The results were presented as a discussion paper and they were asked whether the findings reflected their clinical practice.

3.2.2 Results

Six focus groups were carried out with all participants making an active contribution. The groups included between three and six participants (twenty-seven in total), they are detailed in Table 3.4.

Table 3.4 Participants in the focus groups

Group	No.	Grade	Type of work
1.	6	Senior I or clinical specialist	1st year MSc students
2.	4	Senior I or clinical specialist	2nd year MSc students
3.	3	2x Senior II, 1x Senior 1	Teaching hospital - neurology service
4.	4	1x Senior II, 2x Senior I, 1x Supt III	Regional rehabilitation unit
5.	4	2x Senior II, 2x Senior 1	DGH acute medical & elder rehab wards
6	6	1xSupt II, 2xSupt III, 3x Senior I	Regional Rehabilitation Unit and district stroke service

No. = number of participants, Supt = superintendent physiotherapist
DGH = district general hospital

After six groups, no new insights were being added to those obtained from previous groups. This indicated that saturation had been reached so data collection ceased. All of the groups went well. The use of the photographs and flip charts generated a lively discussion and many items were given to be listed. There was focused and animated discussion with participants agreeing or disagreeing with suggestions and wider discussion developed on why those items were important and how they could be measured. Generating discussion was not a problem. all participants contributed and no individual group members dominated. One group (group 5) was less forthcoming than the others and needed more prompting from the facilitator. It was subsequently discovered that there were strains in the working relationship between some of the participants. However the content of the data from this group did not differ from the other groups. The data from the focus groups is shown in Appendix I

3.2.2.1 General Points

None of the participants used formal assessment models or objective measurements for posture and balance. Three units and some of the MSc students routinely used general outcome measures (such as the Functional Independence Measure) but these were not considered specific enough to inform or assess physiotherapy treatment.

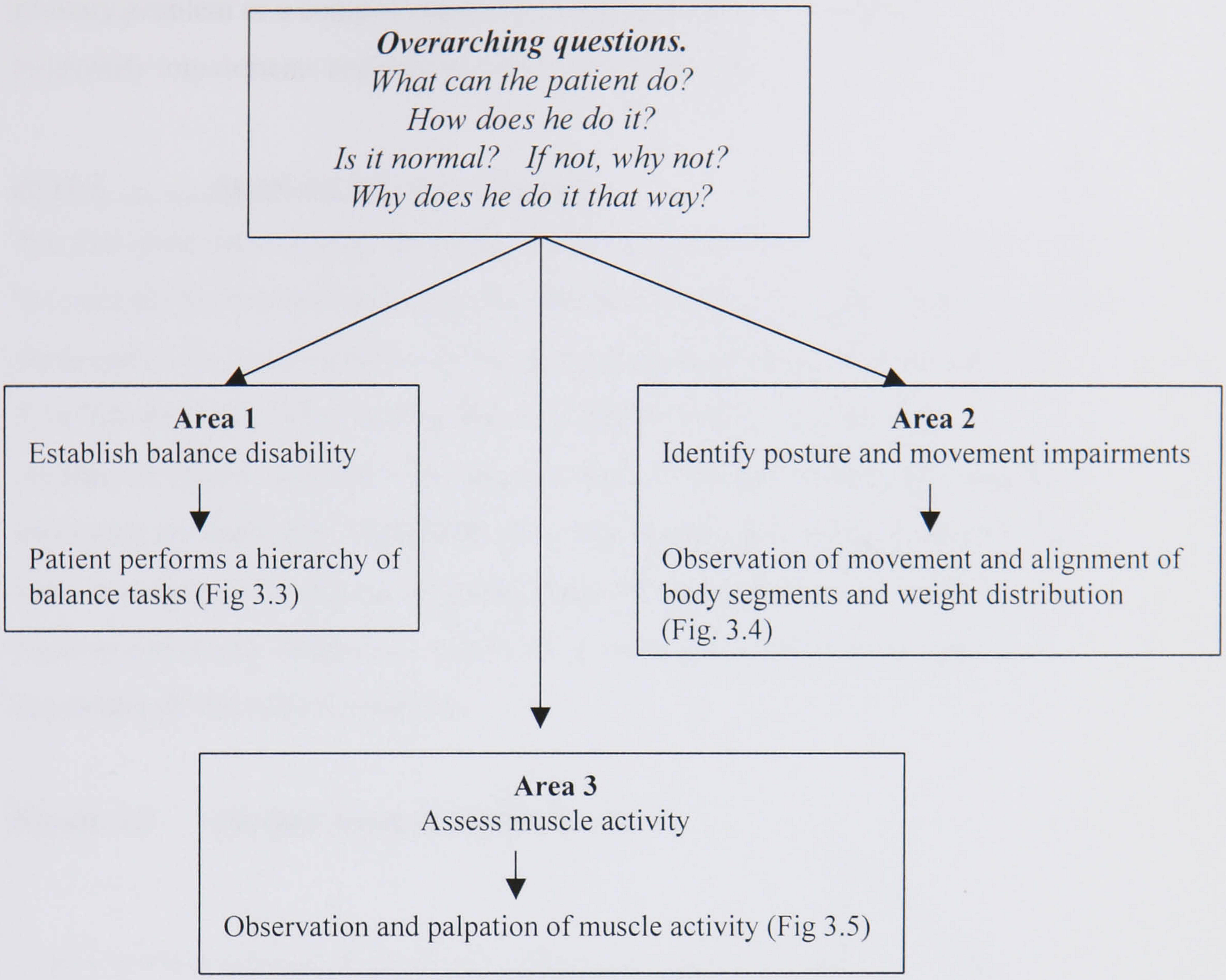
The participants discussed the question of representativeness - whether the photo of the `typical' patient could capture information about the assessment of all patients. If another photo of someone with a more or less severe hemiplegia had been shown would the data be different? There was consensus that the emphasis may be on different features with different patients but the process and the range of features considered would be the same. For instance if someone had a more severe hemiplegia, the emphasis may be on the posture of their head and static sitting balance, rather than posture of their trunk and pelvis. For a milder hemiplegic, the emphasis may be on the posture of their pelvis and leg in standing and stepping activities.

Despite the specific question asked at the beginning of the session, the discussion often ranged broadly and most of the therapists felt that the whole process of assessment and clinical reasoning could not be separated from what was assessed. The detail of the data is considered according to the categories outlined in Figure 3.1

3.2.2.2. The Assessment Process

The clinical reasoning during the assessment process was complex (Figure 3.2) and there was no set structure or order to the process. Different aspects of assessment were often tested at the same time, the order of testing and the emphasis given to different aspects of assessment depended on the individual patients' problems revealed during the assessment. Much of the decision-making was implicit but it was possible to tease out the underlying structure of the assessment process. The overall purpose of the assessment was to establish the patient's abilities - what s/he could or could not do, but more information about how s/he moved was needed to be able to choose the appropriate goals and treatment techniques. During the assessment the physiotherapist was seeking the answers to several over-arching questions: What can the patient do? How does s/he do it? Is it normal? If not, why not? Why does s/he do it that way? The overall process is summarised in Figure 3.2 and each area (1-3) is described in more detail below. The data from the focus groups is shown in Appendix I.

Figure 3.2. The clinical model for assessment of posture and balance in people with stroke.

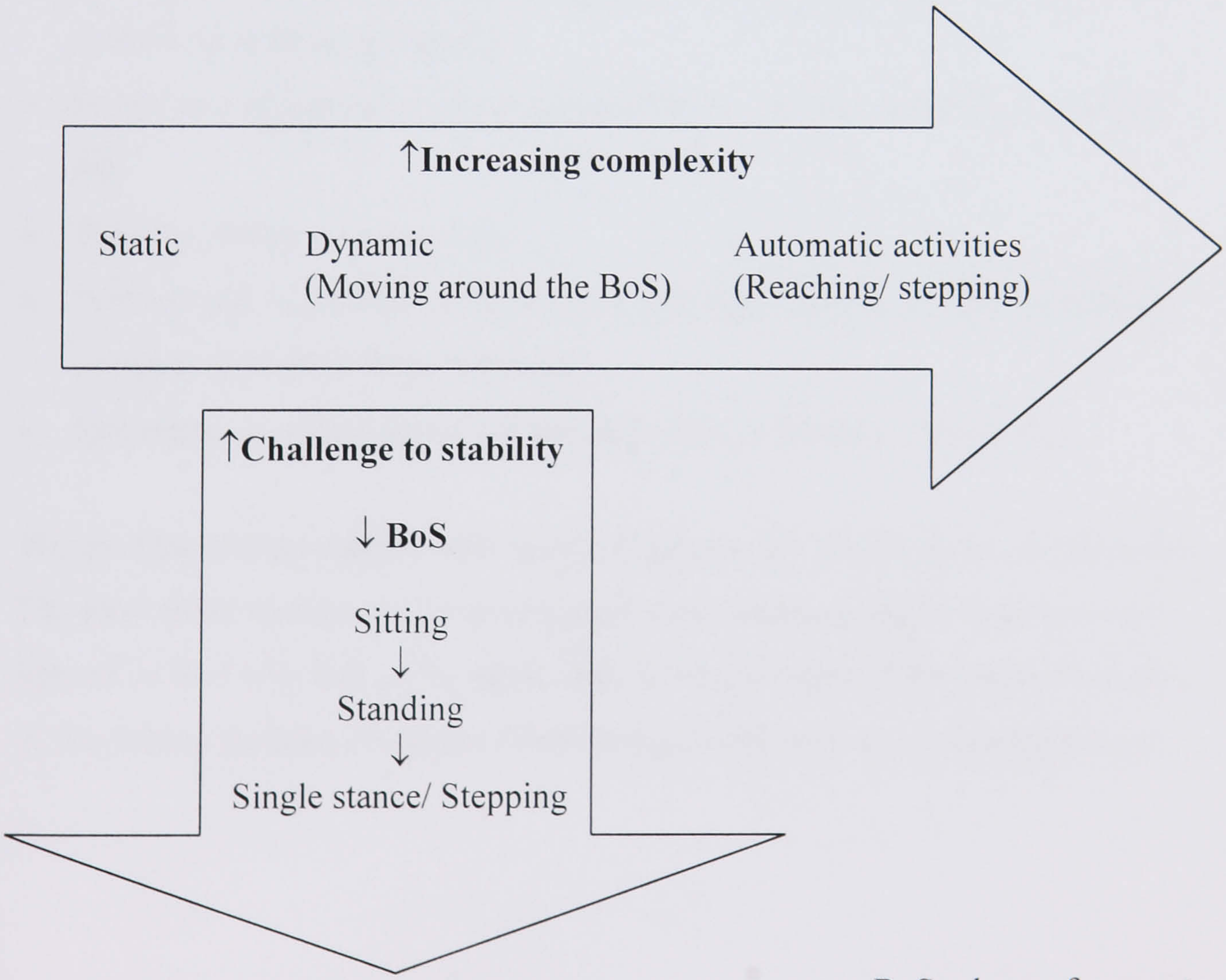


Three areas of assessment were identified. Area 1 addressed what the patient could do by assessing balance disability. This was done by observing the patients ability to perform a series of increasingly demanding balance tasks. The information was used to establish their functional abilities which was used to inform treatment techniques and the assistance they required in every-day activities. Area 2 addressed how the patient performed the balance tasks and whether it was normal by assessing the impairment or abnormalities of posture and movement. This was done by observing the alignment and movement of body segments, particularly of the head, trunk and pelvis as the patient performed the balance tasks. Area 3 assessed muscle activity to establish why the patient performed the tasks in the way that he did. This was done by observation and palpation. The physiotherapists decided whether the activity was increased or decreased or showed altered muscle length, and whether any abnormality was a primary problem or a compensation. The information from area 2 and 3 was used to identify impairments and inform treatment techniques.

Area 1 Assessing balance disability

The first issue the physiotherapists needed information about was an indication of the patient's functional abilities and limitations. To define this, the therapists asked the patient to perform a number of increasingly demanding balance tasks (Figure 3.3). The demands of the balance task were progressed by increasing the complexity of task (moving from static to dynamic to automatic movements) and increasing the challenges to stability. This was done by decreasing the size of the base of support of the position (moving from sitting, to standing, to stepping and walking activities). These tasks were done actively, passively or with assistance depending on the patient's abilities.

Figure 3.3 The hierarchy of balance tasks.



BoS = base of support

<u>Area 2 Assessing posture and movement</u>

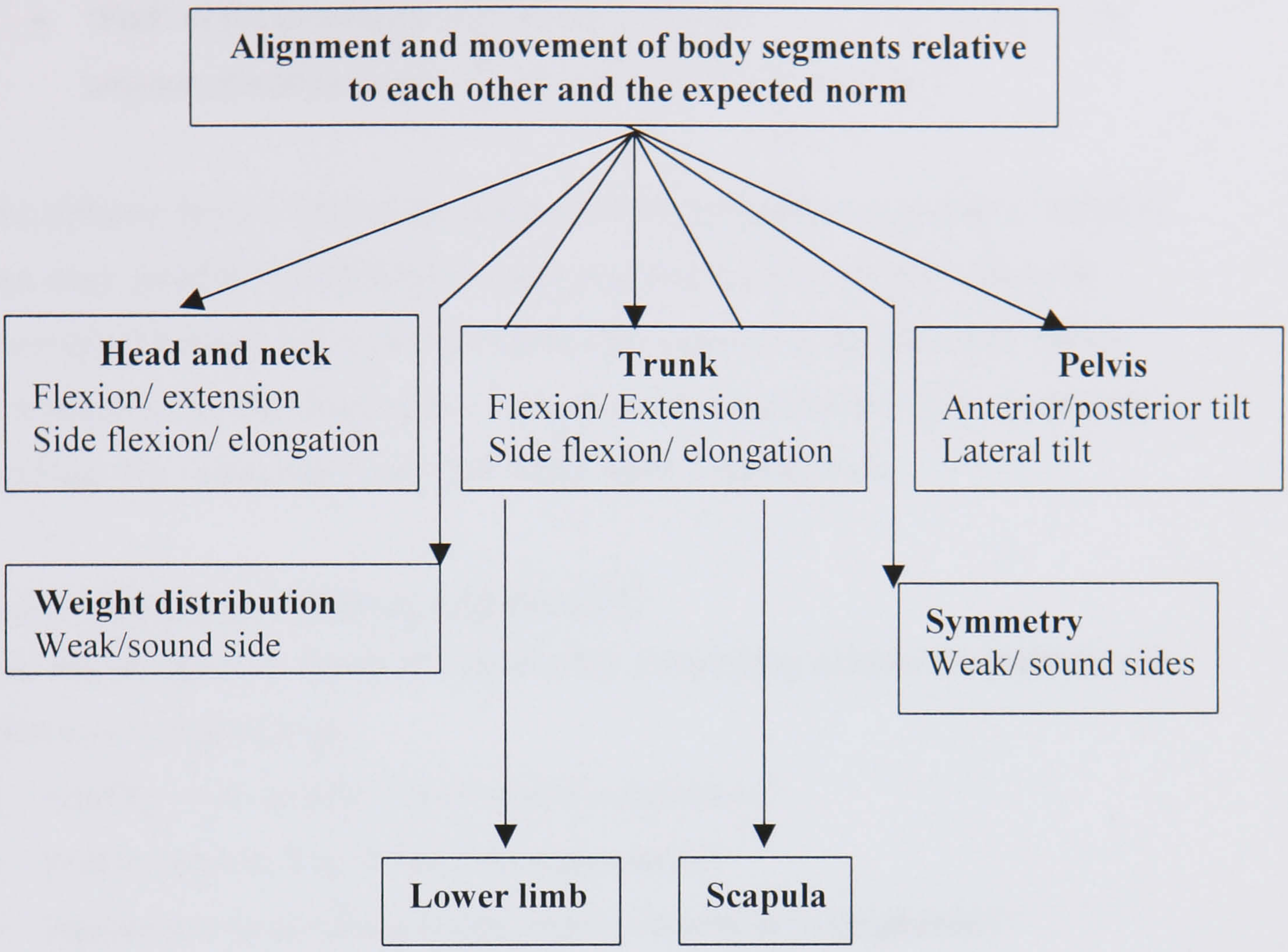
As the patient performed the balance tasks, the therapist said they identified impairments that prevented or limited the patient's function by considering weight distribution and the alignment and movement of body segments, relative to each other and relative to "what they would expect for that patient" (the expected norm). The therapist's view of "what they would expect for that patient" was based on personal experience. If the patient was unable to perform a task normally, the therapist would give physical support and guidance. This was to establish a) whether s/he was able to perform with assistance and b) how much assistance was required, which gave an indication of the severity of the disability.

All groups defined alignment as 'the position of body segments (or joints) relative to each other'. The alignment and movement of most body segments in most planes was considered (Figure 3.4):

- Position of the Head and Neck – flexion / extension, side flexion (including head righting response)
- Position or Alignment of Trunk – flexion/ extension, side flexion/ elongation (including righting reactions),
- Position or Alignment of the Pelvis and Hips – anterior/ posterior tilt, lateral tilt
- Position of hips, knees and feet
- Position and Alignment of the scapula, position of the upper limb (including distance of the limb from the trunk)
- Symmetry – comparison of left and right side, or sound and weak side

Weight Distribution was the other aspect of posture the physiotherapists assessed. The most usual method used was to compare the weight going through the weak buttock or foot with that of the sound side, but the position of the centre of gravity (CoG) within the base of support (BoS) and postural sway were also mentioned.

Figure 3.4. Assessment of posture and movement



<u>**Area 3 Assessing muscle activity**</u>

As the abnormalities of posture and movement were identified the physiotherapists were also considering <u>why</u> these were occurring. Was the muscle activity abnormal, if so how did it differ form normal?

Three types of altered muscle activity were identified (Fig 3.5):

- Decreased muscle activity (flaccidity, low tone, weakness)
- Increased muscle activity (increased tone, hypertonicity, spasticity, muscle imbalance)
- Muscle or soft tissue shortening (contracture, loss of range).

Altered muscle activity was always assessed subjectively. Several different methods were identified:

- Observation of alignment of body segments
- Observation of muscle activity

- Palpation
- Resistance to movement and stretch
- 'Feel' of the movement - the amount of resistance to a movement or the amount of assistance the patient required to perform a task.

The different types of altered muscle activity did not happen in isolation. In fact it was quite possible that all three (increased or decreased activity or soft tissue shortening) could occur around the same body segment at the same time. When this occurred the next stage was to identify which abnormalities were the primary problem (the underlying cause) and which were compensations.

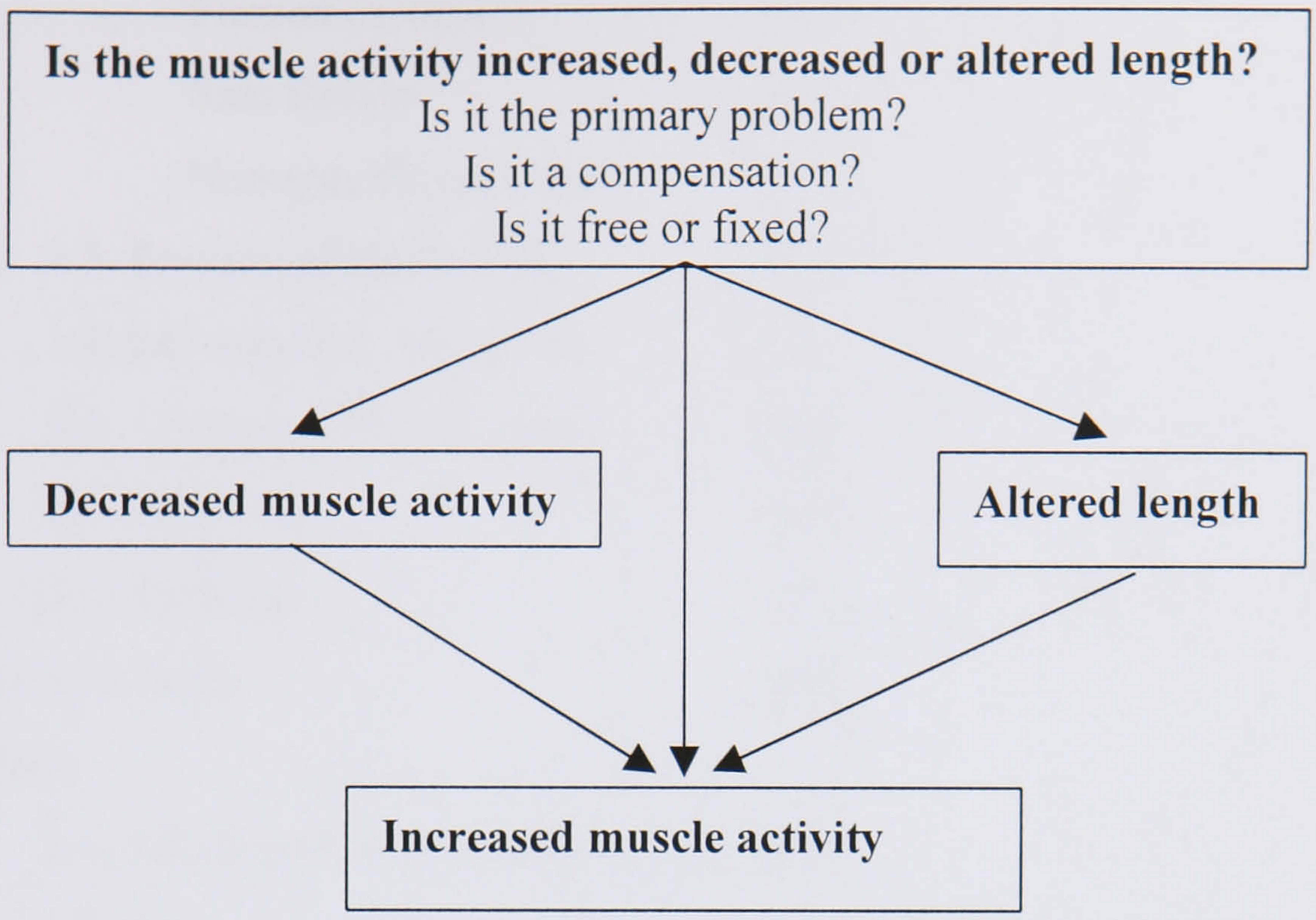
Is the abnormality primary or compensation?

The following points were considered when establishing whether the problem was primary or compensatory.

- Activity – was muscle tone increased or decreased?
- Was the activity free or fixed or compensation?
- Was muscle or soft tissue length normal, shortened or lengthened?

When increased muscle activity or muscle or soft tissue shortening was noted the therapists questioned whether this was the sole cause of the abnormality or whether it was to compensate for an underlying weakness elsewhere. For instance, a weakness in one muscle group may be compensated by increased activity in other muscle groups around the same body segment or in adjoining body segments in order to maintain stability. If the activity was constant to stabilise a joint or to maintain a position, it was referred to as 'fixing'. Whether the muscle activity could be released to allow movement of the joint was important, as it would effect the movement present at the joint and so function. This was referred to as 'free or fixed?' If the muscle activity could be released it was 'free', if it could not be released it was 'fixed'. Compensation was considered similar but different to 'fixing'. Compensation may be used to produce a movement rather than to give stability, which was the characteristic of fixing.

Figure 3.5 Assessing altered muscle activity



Establishing which abnormal muscle activity was the primary problem and identifying compensations was considered a complex topic. The physiotherapists felt that patients varied in the way in which primary problems and compensations manifested themselves, so that it was not possible to stereotype the expected muscle activity or abnormalities. Patients not only varied as individuals but also changed as their recovery progressed.

3.2.2.3 Analysis

Reliability of Categorisation

There was a high degree of agreement about the categorisation. Two assessors felt that one item should be allocated to a different sub-category. One assessor felt that two items in the miscellaneous section could be allocated elsewhere but as the other assessors disagreed no change was made. The percentage of agreement for each sub-category is indicated below.

1. Purpose	100%
2. Process and clinical reasoning	100%
3. Alignment	
3.1. Position of Head and neck	100%

3.2. Position or Alignment of trunk	
Flexion/extension	100%
Side flexion	95%
Non-specific or mixed	100%
3.3. Position of Pelvis/ Hips	100%
3.4. Leg and Foot Alignment	100%
3.5. Position of Scapula and UL	100%
3.6. Symmetry	100%
4. Weight distribution	100%
5. Muscle Activity	100%
6. Others	
6.1. Sensation/ Perception/ Cognition	100%
6.2. Face	100%
6.3. Miscellaneous	84% (1 assessor disagreed with 2 items)

Validity of Analysis

Internal validity was assessed by presenting the results (all of Section 3.2.2) to two of the original groups (Groups 2 and 5) and asking them whether the results a) reflected their clinical practice and b) whether anything should have been included that had been omitted. The participants agreed with the findings and no changes were made. This process was done by informal discussion and negotiation. No formal record was made of the meetings.

External validity was checked by presenting the results to invited participants at a conference on outcome measures in rehabilitation units (as described in Section 3.2.1). Participants were asked whether the results a) reflected their clinical practice and b) whether anything should have been included that had been omitted. The session was organised as a general discussion session and the contributions were not formally recorded other than brief notes of any decisions. These were made by the facilitator during the session. There was some discussion about whether pain and sensation/proprioception should be included as physical features, but it was concluded that that these were additional features and not central to the assessment of posture and balance.

3.2.3 Discussion

3.2.3.1 The method

The method developed to address the aim of this study differed from a conventional focus group methodology in several ways. Firstly, the information required was a factual report of what physiotherapist said they did rather than an exploration of their opinion, knowledge or understanding. Consequently the discourse and interaction between individuals, and differences and similarities between groups was not an issue, so the data collection and subsequent analysis did not include these issues. These differences in purpose, data collection and analysis may lead to the question whether the method used was actually a focus group and whether it might have been more accurately called a workshop or informal discussion group. However, the most important aspect of any study design is that it addresses the aim of the study. The method developed here, be it a focus group or workshop, proved very effective in obtaining a snapshot of physiotherapists' stated practice.

The groups ran smoothly and all participants actively contributed to the discussion. Each group included between three and six participants (twenty-seven in total). Groups of between six and twelve participants are usually recommended, and three is generally considered a minimum number (Morgan 1988, Kitzinger 1995). However in this case, these numbers were considered satisfactory as each group generated a lively discussion. Larger groups may have reduced the level of involvement by some members, and would have made it difficult to ensure that all the suggested features were heard and listed.

Reliability and validity of qualitative data is always difficult to establish, however a number of techniques have been included in this study to address these issues. Firstly, focus groups were held until saturation was obtained which meant that the author could be reasonably confident that all of the important features had been included and the content of the data was similar for all the groups, indicating some degree of reliability and validity. Three independent assessors checked the categorisation of the features and a high degree of agreement was found, indicating reliability and giving some validity to the way that the clinical model had been drawn from the data.

Internal validity was established by presenting the findings to groups of original participants. External validity was established by presenting the results to a group of physiotherapists who had not previously been involved in the study. In both cases participants agreed with the results as a representation of their clinical practice. The use of a group setting to check the validity may have inhibited dissenting voices. Seeking individual views of the validity and assessing this more formally, perhaps through a questionnaire would have given a more robust indication of the extent of agreement or disagreement with the results.

The selection of participants inevitably had an element of convenience as they were drawn from people known by the author and who were local to Brunel University. They did however cover a wide spectrum of types of clinical unit, and of experience and so were a reasonably representative group of neurological physiotherapists. The second group of physiotherapists (who participated in the external validity testing) were drawn from all areas of the UK and so could be considered representative.

The questions for the discussion were very specific in order to focus the participants' attention as specific, factual information was sought. In hindsight, less specific questions may have elicited richer data about the reasoning process, but this would have detracted from the information gained about the physical features. More focus groups could have been conducted asking specifically about the reasoning process, either by returning to the participants or by recruiting new groups, but this was not the purpose of the study. It would also have increased the commitment required from the participants which may have made them reluctant to participate, and would have increased the time and financial resources required which may have become prohibitive.

3.2.3.2 The model

The lack of explicit descriptions and definitions of clinical practice have limited the development of evidence-based practice for many years (Shepherd 1991; Partridge 1994; Ashburn 1995). Although there have been attempts to describe the content of therapy by `unpacking the black box' (Ballinger et al 1999; Pomeroy et al 2001), there has been very little attempt to explain the clinical reasoning process behind the choice of therapy. This model is one of the few accounts of a clinical reasoning process in neurological physiotherapy. Although the process was complex and lacked a pre-defined structure and much of the clinical reasoning implicit, physiotherapists were able to articulate their thought processes and to deduce an explicit model.

The practice of neurological physiotherapy appears to vary in different geographical regions and between different types of clinical service (especially neurological or elder care services) (Davidson & Waters 1999; Lennon & Ashburn 2000), so it may be that this model only represents a fraction of physiotherapy. However, the participants came from several different services (medical and elder care units, acute and rehabilitation services, specialist and district level) and all felt that the model represented their practice. The participants in the external validity group were drawn from all areas of the UK and again they felt that the model represented their practice, so it is felt that the model has overcome these variations. However further research with more formal assessments of participants' opinion of the model is needed to confirm this. Any model of practice needs to reflect current practice, and this is known to change and develop with time (Lennon & Ashburn 2000), so physiotherapists' practice would need to monitored periodically to ensure that the model still reflected current practice and be adapted if necessary.

The methodology and basic model developed here could be used to develop models for assessment of other aspects of physiotherapy such as mobility and upper limb function. Although the `important physical features' would differ with other aspects of physiotherapy, the actual clinical reasoning process may be essentially the same. The methodology could also be used to develop models of treatment.

3.2.3.3 Comparing the experiential literature and physiotherapists' practice.

There was much agreement between the physiotherapists' actual practice and the experiential literature but also some differences. Unlike the experiential literature, the physiotherapists were able to articulate a clinical reasoning process describing what they did when carrying out an assessment, how they did it and what the information was used for. This was a complex task but the resulting model was accepted as representing clinical practice.

The physical features the physiotherapists considered while assessing P&B was similar to those highlighted in the experiential literature. Using subjective observational analysis and comparison with expectations of 'normal' were constant themes for both the clinical physiotherapists and the textbooks. Equally, the importance of alignment and movement of body segments, particularly the head, trunk and pelvis was consistent, as was the importance of muscle activity. The hierarchy of balance tasks the physiotherapists used to identify the patients functional balance abilities were similar to that suggested by Carr &Shepherd (1987).

It was not the specific purpose of this study to compare and contrast the experiential literature and actual practice but an important difference: The importance of abnormal muscle activity did emerge. Many of textbooks saw increased muscle activity as the primary problem (Bobath 1990; Davies 1990, 1994; Edwards 1996; Johnstone 1987) and the cause of other motor impairments and limited function consequently they saw muscle activity as the issue for analysis for its own sake. In contrast, the physiotherapists appeared to attach more importance to the assessment of functional balance abilities and saw the assessment of motor impairments as a way to identify what prevented or limited function rather than a means in itself. They tended to see increased muscle activity as a compensation for another underlying problem (decreased muscle activity or altered muscle length) rather than the main problem. This could represent an important change in physiotherapists' reasoning which merits further investigation.

3.3 Conclusions to Chapter 3

The aim of this chapter was to examine how physiotherapists assessed P&B in order to inform the content of a new outcome measure. This has been achieved albeit by using an unusual methodology. The process of assessment has been examined in detail with clinical physiotherapists who were able to articulate a clear, though complex clinical reasoning process. This has been combined with the tasks that patients performed and the physical features the physiotherapists noted to form a model of clinical assessment.

The results of this study have identified the essential features that need to be included in an assessment of posture and balance if it is to inform clinical practice. There was general consensus between the physiotherapists and the experiential texts about the important physical features to be noted during assessment. These were:

a) the alignment and movement of body segments particularly the head, trunk and pelvis
b) weight distribution
c) muscle activity.

These features were assessed subjectively, mainly by observation.

The information from this Chapter now needs to be combined with the results of the systematic review (Chapter 2) that identified ways in which P&B may be measured to develop a prototype outcome measure.

Chapter 4

Developing the Brunel Balance Assessment

Introduction

Despite extensive investigation of the published research literature, experiential literature and current clinical practice, no methods of assessing posture and balance that met all the utility criteria and reflected clinical practice have been identified. In Chapter 3, the essential content of an assessment of balance disability and posture/movement impairment was identified from clinical physiotherapists and the experiential texts. The results indicated that a measure of balance disability should test the patient's ability to perform a series of increasingly demanding tasks. Posture and movement impairment should be measured by the alignment and movement of body segments, particularly the trunk and pelvis, weight distribution and symmetry. In Chapter 2, a number of possible measurement methods were noted. Hierarchical ordinal scales fulfilled all the utility criteria except that sensitivity to short-term change. Functional performance tests met all the criteria except they were not suitable for a wide range of disabilities. However, it may be possible to combine the best aspects of these two methods to develop a measure of balance disability. The use of inclinometers had possibilities as a measure of alignment and movement of body segments and simple bathroom scales may be suitable to assess weight distribution.

In this Chapter the development of a new assessment of posture and balance using the structure developed in Chapter 3 and the possible methods identified in Chapter 2 will be described. The new assessment needs to meet the original utility criteria and overcome the problems identified with other measures. These were to be:

- reliable
- standardised
- valid as a measure of balance disability and posture/ movement impairment
- suitable for a wide range of severity of hemiplegia, including people who need assistance and support
- sensitive enough to detect short-term change
- portable, simple, cheap and quick to use

- suitable for clinical and community settings
- relevant to and reflective of physiotherapy practice and the functional movements of every day life.

SECTION 4.1 DEVELOPING THE NEW ASSESSMENT

4.1.1 The Balance Assessment

From the systematic review of the literature (Chapter 2), a combination of two assessment methods (a hierarchical ordinal scale and functional performance tests) was identified as a possible way to assess balance with the advantages of both methods but avoiding their disadvantages. A hierarchical ordinal scale could be developed to test the patient's level of balance ability. For each level of the scale there could be a functional performance test. This would provide a more sensitive and accurate measure of performance. The scale could be used to measure the long-term changes such as those due to rehabilitation, while the functional performance tests would be used to measure shorter-term changes.

The results of the focus groups in Chapter 3, revealed how the clinical physiotherapists assessed a patient's balance by testing their ability to perform a series of increasingly demanding tasks. There were two elements which increased the demand made of the patient; i) the tasks became increasingly complex - progressing from static to dynamic movements and ii) and increasing the challenges to stability by making the base of support smaller - moving from sitting to standing to stepping activities. The new assessment would need to reflect this hierarchy.

There was a mismatch between the terms used in clinical practice and that used in the research literature and so a first step in developing a hierarchy of balance tasks was to clarify the terms and definitions used. When clinicians referred to 'static balance' they referred to activities in which the patient maintained a static position. However this ability is categorised in more detail in the research literature, where three different levels are used:

- assisted static balance - when another person gives assistance to maintain the position (Fugl-Meyer et al 1975; Carr et al 1985; Partridge et al 1987; Collen et al 1991; Berg et al 1992a; Gowland et al 1993)
- supported static balance - when the patient uses upper limb support to maintain the position (Fugl-Meyer et al 1975; Carr et al 1985; Collen et al 1991; Berg et al 1992a)
- self-generated movements which require maintenance of a position while the patient moves another body segment (Carr et al 1985; Duncan et al 1990; Goldie et al 1990; Berg et al 1992a; Hill et al 1996; Carr & Shepherd 1998).

Independent balance infers that the person does not require assistance or support. The clinicians used another term that does not occur in the research literature - dynamic balance. This refers to the patient's ability to maintain their balance while moving within the base of support and to the limits of stability (such as the functional reach). It may also include changing the base of support (such as stepping). These levels of dynamic balance are self-generated movements. The research literature (particularly those based on the Brunnstrom approach) also included responses to external perturbations (e.g. Fugl-Meyer et al 1975; Lindmark & Hamarin 1988) but this was not used by the British physiotherapists.

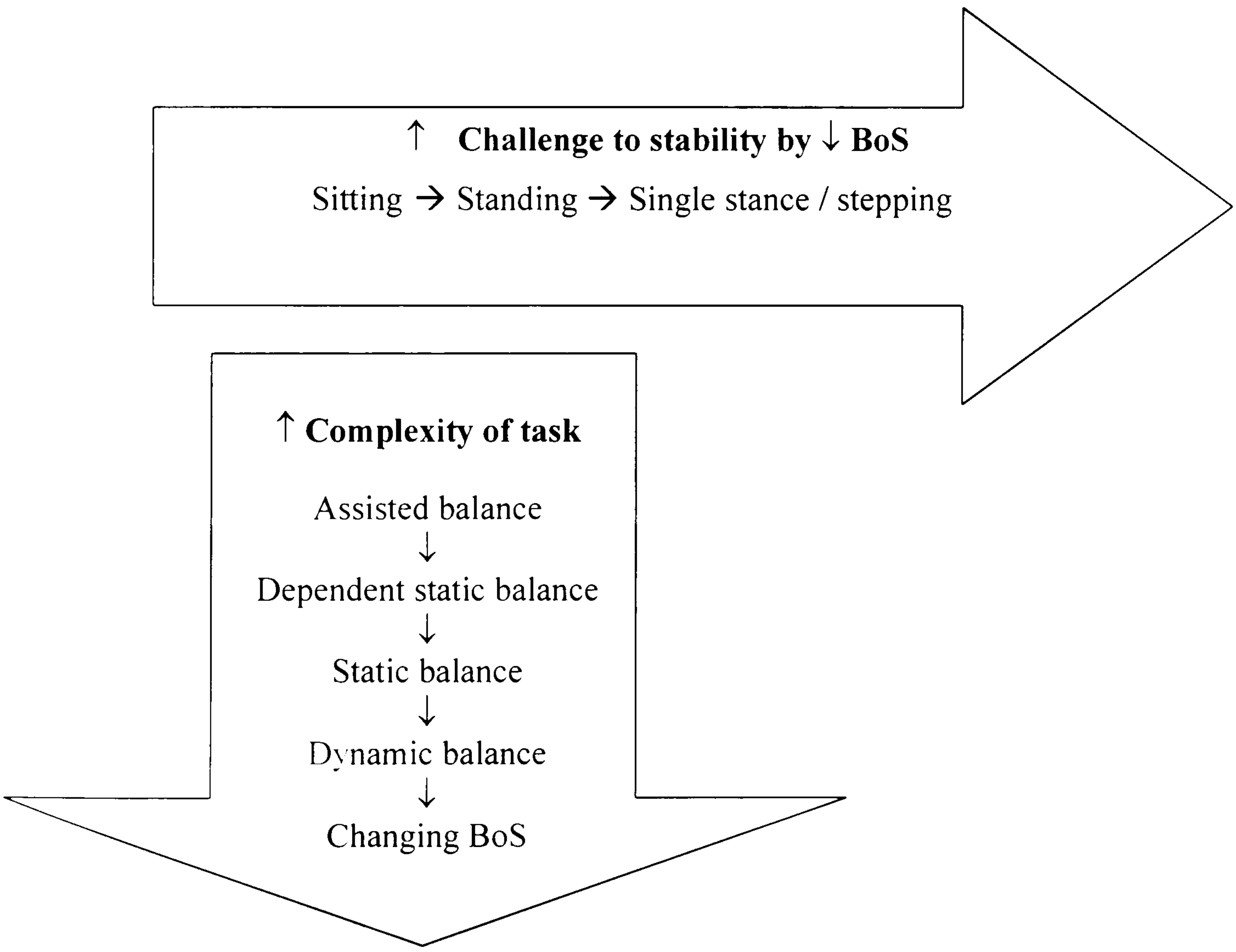
From this mixture of terms, four stages of balance were defined, which were used as the basis for the new balance assessment.

- Assisted balance - ability to maintain a static position with some help (either from another person or by using upper limb support)
- Independent balance - ability to maintain a static position without help
- Static balance - ability to maintain a position while performing self-generated movements (moving another body part)
- Dynamic balance - moving within the base of support to the limits of stability using self-generated movements

- Changing the base of support - moving from one position to another. or from one foot to the other.

Both the clinical physiotherapists and the research literature used three levels of balance ability - sitting, standing and single stance/ stepping. This combined with the balance levels (above) would be used as a matrix for the new ordinal scale (Figure 4.1).

Figure 4.1 Matrix to define the hierarchy of balance tasks



N.B. BoS = Base of support

Having developed a matrix to define the hierarchy of balance tasks, methods for measuring each level needed to be identified. The different methods of assessing balance identified from the ordinal scales and functional performance tests in the literature review (Chapter 2) were assessed according to the utility criteria. The possible methods were discussed with two other lecturers of neurological physiotherapy at Brunel University (subsequently referred to as the panel) and decisions were made about the most suitable tests to use. The decisions drew on the panel's collective clinical experience and our impressions from informal feasibility testing amongst ourselves.

4.1.1.1 Tests of assisted balance (Table 4.1):

There are two types of assisted static balance. When another person gives assistance, or when assistance is given by upper limb support - holding on (Fugl-Meyer et al 1975; Carr et al 1985; Partridge et al 1987; Collen et al 1991; Berg et al 1992a; Gowland et al 1993). In both cases, the ability to achieve this level of balance was on a pass/fail basis. Assistance by upper limb support fulfilled all of the utility criteria while assistance from another person was difficult to standardise (Table 4.1). So assistance from another person was rejected in favour of use of upper limb support as a test of assisted balance. Then the question of how this should be tested arose. The subjects' ability (or inability) to maintain their balance for a pre-defined period would be tested. Thirty seconds was chosen for the testing period as it was considered long enough to be functionally relevant but not so long that fatigue or boredom would affect performance. The subject's posture was not defined because the important factor was whether they could perform the test rather than the way in which they did it. Any safe posture would be acceptable, but to improve standardisation the subject would have their feet flat on the floor in sitting and hip width apart in standing. ' Hip width apart' was defined as standing with the feet level and positioned so that the medial borders of the feet were a comparable distance apart to the distance between the iliac crests.

Table 4.1 How tests of assisted static balance compared with the utility criteria

Utility criteria	Upper limb support	Help of 1 person
Simple & quick	✓	✓
Equipment	✓	✓
Suitable for different settings		
Research	✓	✓
Hospital	✓	✓
Community	✓	✓
Functionally relevant	✓	✓
Standardisable	✓	✗

✓ = pass this criterion, ✗ = fail this criterion.

4.1.1.2 Independent balance (Table 4.2)

In previous outcome measures, tests of independent balance all assessed whether the patient could maintain the position without upper limb support or assistance of another person (Fugl-Meyer et al 1975; Carr et al 1985; Collen et al 1991; Berg et al 1992a). There was however a lot of variation in how this was measured and the position in which the subject was tested (Table 4.2). Either the time the person could maintain balance was measured, or the ability to maintain their balance for pre-defined period was tested on a pass/fail basis; the pre-defined period varied.

The other source of variability was the position in which the person was tested. Sources of variation were; whether the subject's eyes were open and closed, the position of the person's feet in standing and the sitting position (particularly whether they 'slumped' or not, and whether the feet were on the floor).

The panel chose the same testing procedures and positions as for the supported balance tests. Whether the person could maintain balance for thirty seconds was tested on a pass/fail basis. Any safe posture would be acceptable, but to improve standardisation the subject would have their feet flat on the floor in sitting and hip width apart in standing. Timing the period the patient could maintain balance initially seemed an attractive option, as this would provide a ratio level measure. But it was rejected as one lecturer had experience of trying to use this method. She found that if the subject started to lose their balance, the tester's attention naturally went to assisting the patient rather than noting the time. This made it unreliable and difficult to standardise.

Table 4.2 How tests of independent static balance tests compared with the utility criteria

Utility criteria	Timed tests*	Eyes open/ closed	Foot position **	Sitting Position ***
Unit of measurement	*Yes/No*	*Time or Yes/No*	*Time or Yes/No*	*Time or Yes/No*
Simple & quick	✓	✓	✓	✓
Equipment	✓	✓	✓	✗ if use "feet off floor"
Suitable for different settings				
Research	✓	✓	✓	✓
Hospital	✓	✓	✓	✓
Community	✓	✓	✗	✓
Functionally relevant	✓ if 30s or 1min ✗ if 10s or 5mins	✗ if eyes closed	✗ if feet tog. or tandem	✗ if use "feet off floor"
Standardisable	✓	✓		✗ if use "not slumped" position

✓ = pass this criteria, ✗ = fail this criteria.
* Possible timed tests were; 10s, 30s, 1 minute, 3 minutes, 5 minutes
** Possible foot positions in standing were: feet together; wide base (feet hip distance apart); tandem standing (heel of the front foot against the toes of the back foot) and stride standing (one foot a stride length in front of the other).
*** Possible sitting positions were: feet on or off the floor, or a slumped or upright posture.

4.1.1.3 Static balance - maintaining a position during self-generated movements (Table 4.3).

There are three tests of self-generated movements (Carr et al 1985; Goldie et al 1990; Berg et al 1992a; Hill et al 1996). One was nominal - whether the person could turn their head to look behind. This had been used in sitting and standing (Carr et al 1985; Berg et al 1992a). The other two were interval level tests – the Arm Raise Test and the Step Test. The Arm Raise test counted the number of times the subject could raise the sound arm in thirty seconds while standing (Goldie et al 1990). The Step Test counted the number of times the subject could lift one leg on and off a block while maintaining single stance on the other leg (Hill et al 1996).

All the tests fulfilled the utility criteria. The Step Test was chosen as a test of single stance and the Arm Raise test was chosen as the test in sitting and standing. Both tests were performed for 15 seconds as the panel found their own arms or legs ached if testing continued for longer. As the object of the test in this study was to test standing balance, the Step Test would be performed with the subject standing on the weak leg and moving the sound leg.

Head turning was rejected as a test by the panel as it was only a nominal level test and the displacement caused by the movement was very small. So it would not be demanding enough for most patients.

Table 4.3 How the tests of self-generated movements compared with the utility criteria

Utility Criteria	Turn head to look behind	Arm Raise Test	Step Test
Unit of measurement	*Yes/No*	*No. of raises*	*No. of steps*
Simple & quick	✓	✓	✓
Equipment	✓	✓	✓
Suitable for different settings			
Research	✓	✓	✓
Hospital	✓	✓	✓
Community	✓	✓	✓
Functionally relevant	✓	✓	✓
Standardisable	✓	✓	✓

✓ = passes this criterion

4.1.1.4 Dynamic balance: Moving within the base of support.

Three possible tests involved the subject moving to the limits of stability within the base of support (Table 4.4 Carr et al 1985; Duncan et al 1990; Goldie et al 1990; Berg et al 1992a). Reaching forward to the limits of stability was already an established test in standing (Duncan et al 1990), which fulfilled all the utility criteria. Using this test in sitting also appeared to fulfil the criteria so it was chosen as the test of dynamic sitting balance.

Another possibility was to use 'reaching to the floor' as a test of dynamic balance. This has been used a nominal test in sitting (Carr et al 1985) and in standing (Berg et al 1992a). It was rejected by the panel as being difficult to standardise. Further more ratio level measurements were preferred to those at nominal level.

Moving to the limit of stability in stepping would mean transferring the weight from one foot to another. A test method has been described previously by Goldie et al (1990). In this test a pressure gauge is set to 50% of body weight and the number of times the patient transfers sufficient weight on to their weak leg within a set time is scored. The equipment is not commercially available in Europe, nor is it suitable for use in the community. So an alternative way of measuring weight transfer from one foot to the other was needed. The panel devised the Weight-Shift Test. In this test the patient stands with the weak foot in front of the sound foot (step standing) and transfers their weight from one foot to the other and back again. The number of times the weight is transferred onto the weak foot in 15s is counted (Full details in Appendix III).

Table 4.4. How the advanced dynamic balance tests compared with the utility criteria

Utility criteria	Reaching		Weight shifts
	Forwards	**To the floor**	
Unit of measurement	Distance or reps	Reps or Y/N	Reps or % body wt
Simple & quick	✓	✓	✓
Equipment	✓	✓	✓
Suitable for different settings:			
Research	✓	✓	✓
Hospital	✓	✓	✓
Community	✓	✓	✓
Functionally relevant	✓	✓	✓
Standardisable	✓	✗	✓

✓ = pass this criterion, ✗ = fail this criterion.

4.1.1.5 Changing the base of support (Table 4.5).

Three tests that involved changing the base of support were found from the literature in Chapter 2 (Table 4.5), all involving standing and stepping tasks. Changing the base of support in sitting was investigated by the panel but rejected because it would involve moving to another position and activity (i.e. sit to stand). In standing, changing the base of support involves stepping activities so it was merged with that section. Three tests were identified. A nominal test or timed test of whether the subject could turn 360° (Berg et al 1992a); walking speed (Wade et al 1987, Collen et al 1990) and step-ups. All the tests fulfilled the utility criteria. The panel chose walking speed and step-ups as tests of changing the base of support. This was based on personal preference and because they offered an obvious progression. 5m was chosen as a distance for the walk test as this would be feasible in most community settings (Collen et al 1990). The panel discussed the use of walking as a test of balance ability. All felt that, in terms of balance, walking involved two tasks - the ability to maintain balance in single stance (during stepping) and the ability to transfer the base of support between double and single stance. As such it could be seen as a single stance task or a changing base of support task, or both. Another complication was the use of walking aids. The panel felt that people who could walk without an aid were functionally and clinically different to people who needed an aid to walk. They thought that both levels of walking should be included because they were different balance tasks. When considered in terms of the balance task involved, walking with an aid could be defined as a supported single stance task as the person could take support from the walking aid when in single stance. Walking without an aid could be defined as changing the base of support between double to single stance. So the final choice of tests was *5m-walk test with an aid* (as a measure of supported single stance), *5m walk test without an aid* (as a measure of changing the base of support between single to double stance) and *step-ups* (as an advanced measure of changing the base of support between levels).

Table 4.5 How the tests of changing the base of support compared with the utility criteria.

Utility criteria	**Walking**	**Turn 360°**	**Step ups**
Unit of measurement	Speed	Y/N	No. of reps
Simple & quick	✓	✓	✓
Equipment	✓	✓	✓
Suitable for different settings:			
Research	✓	✓	✓
Hospital	✓	✓	✓
Community	✓	✓	✓
Functionally relevant	✓	✓	✓
Standardisable	✓	✓	✓

✓ = pass this criterion

4.1.1.6 Constructing the New Balance Assessment

Having identified suitable tests for the balance hierarchy, the proposed tests were entered onto the balance test matrix hierarchy (Table 4.6).

The content of the matrix was converted into a scale (referred to as the Brunel Balance Assessment, Table 4.7) and this was used for subsequent testing and development as an outcome measure. The way in which the tests were performed is described in the instruction manual (Appendix III). For clarity the Step test was subsequently referred to in the text as the Tap Test to differentiate it from the Step-Up Test.

Table 4.6 The balance hierarchy matrix with corresponding tests.

	Sitting	Standing	Stepping
Supported balance	Nominal test over 30s	Nominal test over 30s	5m walk test with an aid (single stance)
Independent balance	Nominal test over 30s	Nominal test over 30s	Nominal test over 30s (double stance)
Static balance	Arm Raise Test	Arm Raise Test	Tap test (single stance)
Dynamic balance	Forward Reach Test	Forward Reach Test	Weight Shift test (double stance)
Changing BoS I	Not appropriate	5 m walk test without an aid (single ↔ double stance)	
Change of BoS II		Step-up test (change in height)	

Table 4.7 The prototype Brunel Balance Assessment.

Level of balance	Performance test
1. Supported sitting balance	Static sitting with arm support
2. Independent sitting balance	Static sitting without arm support
3. Static sitting balance	Arm raise test*
4. Dynamic sitting balance	Forward reach test*
5. Supported standing balance	Static standing with arm support
6. Independent standing balance	Static standing without arm support
7. Static standing balance	Arm raise test
8. Dynamic standing balance	Forward reach test*

9.	Independent static double stance	Static stride standing without arm support
10.	Supported single stance	5m walk test with an aid
11.	Dynamic double stance	Weight-shift test*
12.	Dynamic single stance	Tap test
13.	Initial change of base of support	5m walk test without an aid
14.	Advanced change of base of support	Step-up test*

* *not previously reported in people with stroke*

4.1.2 The posture assessment

The results of Chapter 3 revealed that physiotherapists assessed posture by observing the alignment and movement of body segments, weight distribution and symmetry while the patient performed the balance tasks. The essential alignments were: anterior-posterior and lateral tilt of the pelvis, flexion-extension and side flexion of the trunk. Methods for measuring these positions and weight distribution were therefore sought for inclusion in the new assessment. Posture and weight distribution was measured at the different levels of balance ability identified in the ordinal balance scale above. The parameters to be measured were therefore:

- Lateral weight distribution in sitting and standing
- Anterior/ posterior weight distribution in sitting and standing
- Lateral tilt of the pelvis during static and dynamic activities (probably at the limits of stability) in sitting and standing
- Anterior-posterior tilt of the pelvis during static and dynamic activities (probably at the limits of stability) in sitting and standing
- Lateral tilt (side flexion) of the trunk during static and dynamic activities (probably at the limits of stability) in sitting and standing
- Anterior-posterior tilt (flexion-extension) of the trunk during static and dynamic activities (probably at the limits of stability) in sitting and standing

In addition, the protocol needed to fulfil the utility criteria identified previously and be suitable for people who needed support and assistance. These conditions were applied to the methods of measuring posture identified in the literature review and are summarised in Table 4.8. Some methods had already been rejected because of they required sophisticated equipment and so were unsuitable for use in the community. These were: movement analysis systems such as Coda or Vicon, platform tests of postural sway, and external perturbation tests (the postural stress test and maximal load test).

From Table 4.8 it can be seen that none of the assessment methods fulfilled the conditions and new methods thus needed to be developed. An inclinometer that could be strapped to the patient and in which the angle reading could be locked (to leave the testers' hands free to assist the patient if necessary) was needed. Despite extensive searching of the research literature and commercial sources no such equipment could be found.

A weight distribution monitor that could be used in sitting with people who could not position themselves easily or accurately was also needed. The use of bathroom scales was a promising initial idea but proved disappointing. Informal pre-pilot feasibility testing showed them unusable, because the weight/loading obtained was very dependent on the position of the subject's bottom on the scales. The smallest of change in position of their bottom, their posture or the position of their feet produced large changes in data. The method could not even be standardised sufficiently to start data collection for pilot testing.

After extensive searching of research and commercial literature no equipment that fulfilled the utility criteria was found. The researcher is now working with the Design for Life Centre of the Dept of Design Engineering at Brunel University to develop such equipment. In discussion with the project supervisor, Prof. De Souza, it was decided that the development of new equipment was beyond the scope (in terms of time and financial resources) of this project and this would continue as a separate project for which funding would be sought. Work would continue to develop the balance assessment with the intention that the eventual posture measures would complement the balance assessment.

Table 4.8 How measures of posture and weight distribution compared with utility criteria and the required measurements

	Force plate platforms	Digital scales	Goniometer	Inclinometer	Distance between bony points	Flexicurve	Spirit level	Plumb line
Trunk : Lateral : Flexion/Extension	N/A	N/A	✓✓	✓✓	✓✓	✓✓	✓✗	✓✗
Pelvis: Lateral Anterior/Posterior	N/A	N/A	✓✗	✓✓	✓✗	✗✗	✓✗	✗✗
Weight Distribution in sitting		Possibly	N/A	N/A	N/A	N/A	N/A	N/A
Lateral	✓							
Anterior/Posterior	✓							
Wt Distribution in standing			N/A	N/A	N/A	N/A	N/A	N/A
Lateral	✓	✓						
Anterior/Posterior	✓	✗						
Simple & Quick	✓	✓	✓	✓	✓	✓	✓	✓
Equipment	✗	✓	✓	✓	✓	✓	✓	✓
Suitable for different settings?								
Research	✓	✓	✓	✓	✓	✓	✓	✓
Hospital	✓	✓	✓	✓	✓	✓	✓	✓
Community	✗	✓	✓	✓	✓	✓	✓	✗
Functionally relevant	✓	✓	✓	✓	✓	✓	✓	✓
Standardisable	✓	✓	✓	✓	✓	✓	✓	✓
Suitable for people who need help?	✓	✗	✗	✗	✗	✗	✗	✗

✓ = passes this criterion, ✗ = fails this criterion N/A = not applicable

SECTION 4.2 TESTING THE BRUNEL BALANCE ASSESSMENT

Having produced a prototype Brunel Balance Assessment (BBA), the next stage was to test the psychometric properties of this outcome measure to establish its utility. This process is described in the following sections. They cover:

4.2.1 Subjects

4.2.2 Scale development in which the appropriateness of the choice and order of items in the ordinal scale was assessed

4.2.3 Responsiveness to change in which whether the assessment could detect change in a patient's performance was tested

4.2.4 Reliability testing in which the stability of a score obtained from the assessment was assessed

4.2.5 Validity testing in which whether the assessment tests the underlying construct in this case, balance was assessed.

4.2.1 Subjects and recruitment

The inclusion criteria for all people with stroke recruited to the studies included were to:

- be over 40 years old
- have had a first-ever stroke causing a hemiplegia which affected mobility and/or balance
- have been fully independent before the stroke
- have no other pathologies which severely limited balance or range of movement in hips or shoulders
- be able to give informed consent.

The desired population was a group representing people with stroke who received physiotherapy. It was important that any balance disability could be isolated to the stroke rather than to other pathologies. Consequently, people with other pathologies that severely affected balance were excluded. Inevitably this excluded many people with stroke who received physiotherapy but it was important that the balance problems tested with the Brunel Balance Assessment were due to stroke in order to define its validity as measure of balance disability post-stroke. Subjects also needed to be able to give informed consent (this was a condition of ethical approval), which also excluded people who would normally receive physiotherapy in the clinical setting. If it was unclear whether the subject would be able to give consent due to language or communication problems, the opinion of the treating speech and language therapist was sought. If the subject had cognitive problems, which might have prevented informed consent being given, medical opinion was sought before the patient was approached to take part.

Subjects were recruited from the physiotherapy departments of several hospitals local to Brunel University: Northwick Park, Hillingdon, Ealing, West Middlesex, and Clayponds Hospitals. Subjects were recruited from the in-patient, out-patient and community services. The physiotherapists treating people with stroke in these hospitals monitored admissions and referrals for people who met the appropriate inclusion criteria. If a patient met the criteria, the physiotherapist would approach them to see if they were interested in taking part in the study. The physiotherapist briefly explained the purpose of the study, what it would involve, and if they were interested, gave them an information leaflet. A non-treating physiotherapist (ST) contacted them at least twenty-four hours later to see whether they wished to take part. If they did, informed consent was obtained and a time and place for the testing was arranged at their convenience. Testing took place wherever the patient received their physiotherapy – at the hospital bedside, treatment area or their home. Ethical approval was obtained from the relevant ethical committees at the Department of Health Studies at Brunel University and each participating NHS trust.

Ninety-two people with stroke were recruited. Data from eighty people (subjects 1-80) were used in the scale development (section 4.2.2). Data from thirty-seven people (subjects 1-37) were used to test responsiveness to change (section 4.2.3) and reliability (section 4.2.4). Data from fifty-five people (subject 38-92) were used to test the validity (section 4.2.5). The allocation of data in the three groups and the demographics details of the groups are shown in Figures 4.2a and b. The inclusion criteria and the recruiting hospitals were the same for all three groups and the allocation of subjects was entirely arbitrary. Reliability and responsiveness to change was tested on the first 37 subjects (subjects 1-37). Then data collection stopped and it was analysed over the Christmas holidays. Having obtained satisfactory reliability, data was collected on a further 43 subjects (subjects 38-80) and the data on the Brunel Balance Assessment from them and the first group was used to test the hierarchical scaling. Data from the second group (subjects 38-80) and another 12 subjects (subjects 80-92) were used to test validity. Data from these 55 subjects (subjects 38-92) were also used in the analysis in Chapter 5.

Figure 4.2a Showing the allocation of subjects to the three data sets.

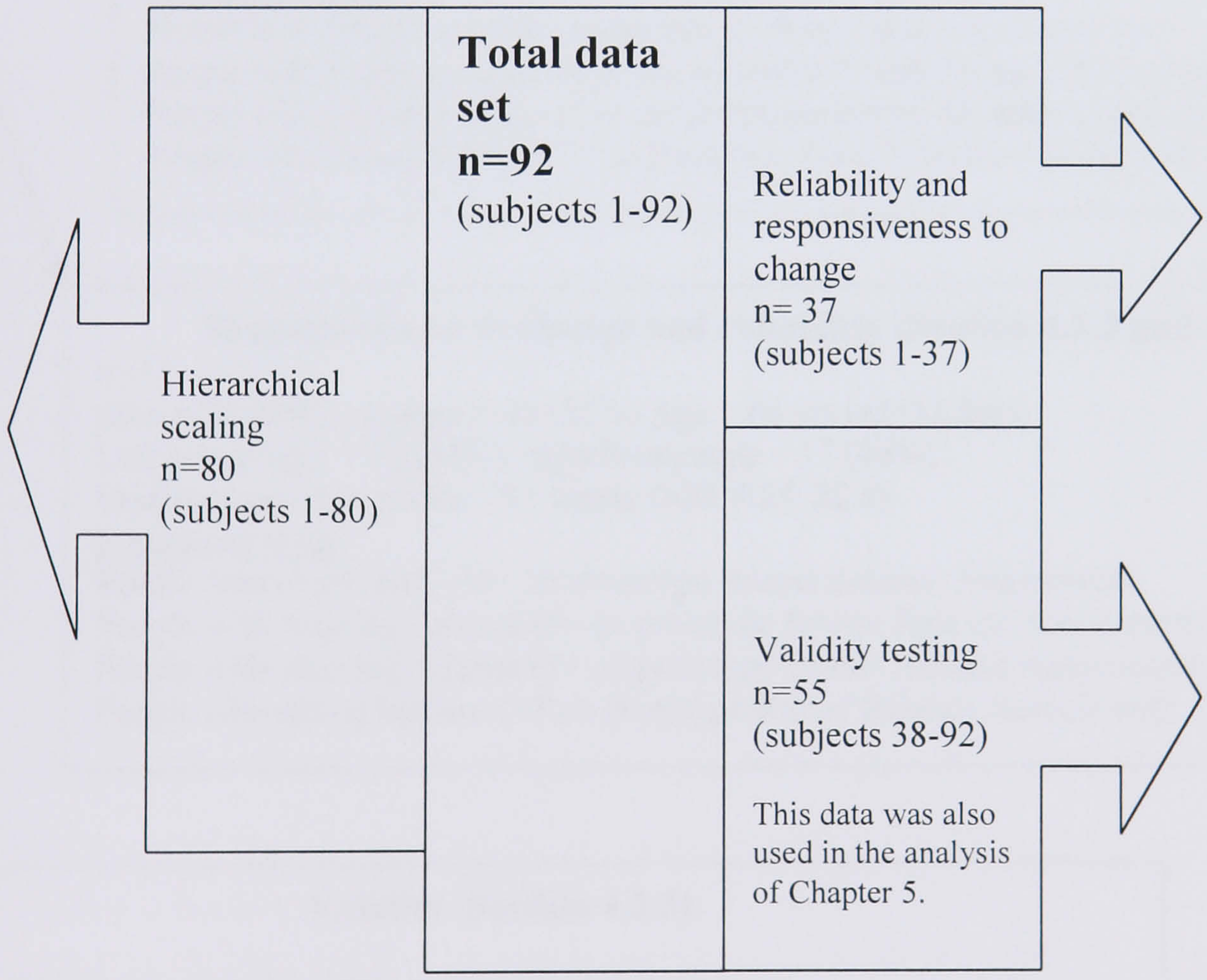


Figure 4.2b The data sets used for testing the Brunel Balance Assessment

Total dataset

n=92
Men = 55 (60%), women = 37 (40%). Age = 67.4 years (sd=12.84, range = 40-88)
Left hemiplegia = 49 (53%), right hemiplegia = 43 (47%)
Median time since stroke = 9 weeks (IQR = 4, 22.25, outliers at 18 months and 8 years)
<u>Functional ability</u>

- People who could walk (scored 10+ on prototype Brunel Balance Assessment) = 41 (44.5%)
- People with stepping balance (scored 8+ on prototype Brunel Balance Assessment) = 43 (47%)
- People with standing balance (scored 5+ on prototype Brunel Balance Assessment) = 27 (29%)
- People with sitting balance (scored <5 on prototype Brunel Balance Assessment) = 22 (24%)

Scale Development (Section 4.2.2)

n=80
Men = 50 (63%), women = 30 (37%) Age = 66.7 yrs (sd=12.7yrs)
Left hemiplegia = 45 (56%), right hemiplegia = 35 (44%)
Median time since stroke = 10.5 weeks (IQR 4.25-22.5)
<u>Functional ability</u>

- People who could walk (10+ on prototype Brunel Balance Assessment) = 37 (46%)
- People with stepping balance (8+ on prototype Brunel Balance Assessment) = 39 (49%)
- People with standing balance (5+ on prototype Brunel Balance Assessment) = 24 (30%)
- People with sitting balance (<5 on prototype Brunel Balance Assessment) = 17 (21%)

Responsiveness to change and reliability (Section 4.2.3 and 4.2.4)

n=37
Men = 24 (65%), women = 13 (35%) Age = 66 yrs (sd=12.8yrs)
Left hemiplegia = 20 (54%), right hemiplegia = 17 (46%)
Median time since stroke = 11 weeks (IQR 4.25, 22.5)
<u>Functional ability</u>

- People who could walk (10+ on prototype Brunel Balance Assessment) = 21 (57%)
- People with stepping balance (8+ on prototype Brunel Balance Assessment) = 23 (62%)
- People with standing balance (5+ on prototype Brunel Balance Assessment) = 8 (22%)
- People with sitting balance (<5 on prototype Brunel Balance Assessment) = 6 (16%)

Validity (Section 4.2.5)

n=55
Men = 32 (58%), women = 23 (42%) Age = 68 yrs (sd=12.8yrs)
Left hemiplegia = 29 (53%), right hemiplegia = 26 (47%)
Median time since stroke = 9 weeks (IQR 4, 24)
<u>Functional ability</u>

- People who could walk (10+ on prototype Brunel Balance Assessment) = 19 (34.5%)
- People with stepping balance (8+ on prototype Brunel Balance Assessment) = 20 (36%)
- People with standing balance (5+ on prototype Brunel Balance Assessment) = 18 (33%)
- People with sitting balance (<5 on prototype Brunel Balance Assessment) = 17 (31%)

SECTION 4.2.2 SCALE DEVELOPMENT

Having chosen items for the prototype scale, the appropriateness of these choices needed to be considered. There are a number of ways that a scale can be constructed. One method is to arrange the items into a hierarchy so the difficulty increases with each item. If the scale forms a true hierarchy, then when a subject fails an item, s/he can be assumed to fail all the following, higher items, or if s/he passes an item it can be assumed that s/he would pass all the lower items (Wade 1992). This means that not all the items need to be tested each time the subject is assessed. Testing can stop once the subject has failed an item, alternatively it can start at a level the subject would find reasonably challenging. This reduces the time and effort required for testing for both tester and subject (Wade 1992). Another advantage of a hierarchical scale is that it gives information about what a patient can or cannot do, rather than how many activities s/he can do (Eakin 1989).

If this limited testing of items is used then it is essential that all the items are homogenous, that is they all test the same basic trait, in this case balance (Barer & Nouri 1989). If the items are homogenous then they should all be moderately correlated to each other and each should correlate to the total score. These two factors form the basis for testing homogeneity or internal consistency (Streiner & Norman 1997). If however, one or more items were very highly correlated to another, such item(s) would not add any new information, and would be unnecessary or redundant (Streiner & Norman 1997). Having redundant items would not only increase the time and effort required for testing but also artificially inflate the internal consistency of the scale, and should be discarded (Streiner & Norman 1997). The hierarchical order, homogeneity and necessity of items are all features of a Guttman-type hierarchical scale (Streiner & Norman 1997) and will be tested in this study.

Method and results

Subjects

Eighty subjects were recruited for the scale development testing as detailed in section 4.2.1. Their balance was tested using the Brunel Balance Assessment as described in Appendix III.

4.2.2.1 Hierarchy

The order of items in the prototype scale was drawn from the theoretical expectation that the subjects' ability to perform the items would decrease as the complexity of the task increased and the stability of the position decreased. If the items were in the correct order the number of subjects passing each item would decrease progressively as the items became more difficult. This was analysed by counting the number of subjects passing each level (Table 4.9, Figure 4.3).

Table 4.9. The pass rates for items of the prototype scale.

Level of balance	Number of people passing the item (%)
1. Supported sitting balance (Sitting with arm support)	80(100)
2. Independent sitting balance (Sitting without arm support)	76 (95)
3. Initial dynamic sitting balance (Arm raise test)	74 (93)
4. Advanced dynamic sitting balance (Forward reach test)	72 (90)
5. Supported standing balance (Standing with arm support)	64 (80)
6. Independent standing balance (Standing without arm support)	56 (70)
7. Initial dynamic standing balance (Arm raise test)	54 (67.5)
8. Advanced dynamic standing balance (Forward reach test)	48 (60)
9. Independent static double stance (Step standing without arm support)	41 (51)
10. Supported single stance (5m walk test with an aid)	37(46)
11. Advanced dynamic double stance (Weight-shift test)	32 (40)
12. Initial dynamic single stance (Tap test)	25 (31)
13. Changing base of support I (5m walk test without an aid)	27 (34)
14. Advanced change of the base of support (Step-up test)	20 (25)

The number passing each item decreased in a step-wise fashion for the sitting and standing items, indicating that the items were in the correct order. but the stepping section was less well ordered. Walking with an aid and the Tap Test were in the wrong order. These were reversed so that the items appeared in order of item difficulty, based on the number of people passing the item (Table 4.10. Figure 4.4), and this order was used in subsequent testing of the scale.

Table 4.10. The revised order of the stepping section of the prototype scale

Stepping Items	No. of people passing the item (%)
9. Independent static double stance (Stride standing without arm support)	41 (51)
10. Supported single stance (5m walk test with an aid)	37 (46)
11. Advanced dynamic double stance (Weight-shift test)	32 (40)
12. Initial change of base of support (5m walk test without an aid)	27 (34)
13. Initial dynamic single stance (Tap test)	25 (31)
14. Advanced change of the base of support (Step-up test)	20 (25)

Assessing the hierarchy of the revised order.

Two tests were used to assess the hierarchy of the scale - or scalability. Firstly, the co-efficient of reproducibility (CR) indicates the degree to which the subject's score is a predictor of his/her response i.e. the likelihood that he would fail all items following the final item he passed and to have passed all the items preceding it. A score of 0.9 or more is considered acceptable (Guttman 1944, Striener & Norman 1997). The Co-efficient of reproducibility (CR) is calculated from the proportion of scaling errors (subjects who do not pass the items in the scale sequence) to the total responses. There were seven scaling errors, which all occurred in the stepping section.

$$CR = 1 - (\text{scaling errors} / \text{no. of items} \times \text{no. of subjects})$$

$$= 1 - (7/14x80) \qquad = 1 - 0.006 \qquad = 0.994$$

The CR will however be influenced by 'extreme subjects' (people who pass or fail all items), or 'extreme items' (items that all subjects pass or fail). Another test, the co-efficient of scalability (CS), takes this into account by assessing the proportion of scaling errors to the number of 'maximum errors', where the maximum errors represent the subjects or items without extreme scores. The CS is always lower than the CR, and so a score of 0.6 or more is considered acceptable (Menzel 1953).

CS (subjects) = 1- (subject scaling errors/max. errors) =1 - (7 / 60) = 1- 0.12 = 0.88

CS (items) = 1- (item scaling errors/max. errors) = 1 -(4/13) = 0.69

Both co-efficients were within acceptable limits indicating that the revised order formed a hierarchical order. The next stage was to assess the homogeneity, or internal consistency of the items.

4.2.2.3 Internal consistency

Two methods were used to assess internal consistency, an item-total correlation and Cronbach's alpha co-efficient. An item-total correlation examines how each item relates to the total score. Any items with a correlation of less 0.2 are judged to be drawn from a different trait to the other items and should be discarded (Streiner & Norman 1997). All the items of the Brunel Balance Assessment had item-total correlations of more than 0.2 and were retained. They are listed below (Table 4.11).

Internal consistency was tested using Cronbach's alpha co-efficient, which is based on all the possible correlations between the items of the scale (Bowling 1997). A score of 0.7 or more indicates acceptable internal consistency (Bowling 1997). The Cronbach's alpha co-efficient was 0.93. This indicated a high degree of internal consistency and that the items of the scale were homogenous.

Table 4.11 The item-total correlation

Test	Item-total correlation
Sitting with arm support	zero variance
Sitting without arm support	0.34
Sitting Arm Raise test	0.43
Sitting Forward Reach test	0.50
Standing with arm support	0.39
Standing without arm support	0.65
Standing Arm Raise test	0.78
Standing Forward Reach test	0.80
Stride standing without arm support	0.81
5m walk test with an aid	0.84
Weight-shift test	0.8
5m walk test without an aid	0.75
Tap test	0.73
Step-up test	0.66

4.2.1.4 Redundancy of items

Scales with a very high level of internal consistency (more than 0.9) may indicate redundancy - items that are essentially reproducing the information from other items. This can be tested using an inter-item correlation, which correlates each item with all the other items in turn. A correlation of more than 0.9 would indicate redundancy and one of those items might be discarded (Streiner & Norman 1995). The inter-item correlation matrix for the Brunel Balance Assessment is shown in Table 4.12. Two correlations indicated some redundancy - between levels 2 and 3, and levels 6 and 7, with scores of 0.89 and 0.94 respectively. This would indicate that subjects that could keep their balance without upper limb support for 30 seconds could also perform at least 2 arm raises and indicates that only one test was required. As the arm raise test is an interval test, and the static test is pass/fail, the static tests were discarded. The revised 12-item scale is shown in Table 4.13

To assess the effect of these changes on the scale, the co-efficient of reproducibility and scalability (for items and subjects) and alpha co-efficient was recalculated for the revised scale. The co-efficients of reproducibility and scalability were identical (0.99 and 0.69 respectively) and the new alpha coefficient was 0.92, which was still acceptable. So the revised scale was used for the next stage of testing, which was the responsiveness to change.

Table 4.13 The revised Brunel Balance Assessment.

Sitting Section	
1. Supported sitting balance	(Sitting with arm support)
2. Static sitting balance	(Arm raise test)
3. Dynamic sitting balance	(Forward reach test)
Standing Section	
4. Supported standing balance	(Standing with arm support)
5. Static standing balance	(Arm raise test)
6. Dynamic standing balance	(Forward reach test)
Stepping Section	
7. Independent double stance	(Step standing without arm support)
8. Supported single stance	(5m walk test with an aid)
9. Dynamic double stance	(Weight-shift test)
10. Change of base of support (double ↔ single stance)	(5m walk test without an aid)
11. Dynamic single stance	(Tap test)
12. Advanced change of the base of support	(Step-up test)

Table 4.12 The inter-item correlation matrix to test for redundancy of items

Items	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Sitting with UL														
2. Sitting without UL														
3. Sitting Arm Raise		**.89**												
4. Sitting Fwd Reach		.69	.77											
5. Standing w. UL		.46	.52	.67										
6. Standing w'o. UL		.35	.39	.51	.76									
7. Standing Arm Raise		.33	.37	.48	.72	**.94**								
8. Standing Fwd Reach		.28	.32	.41	.61	.80	.85							
9. Step Standing		.24	.26	.34	.51	.67	.71	.84						
10. Walking w. aid		.21	.24	.31	.46	.61	.64	.76	.85					
11. Wt shifts		.19	.21	.27	.41	.53	.57	.67	.80	.83				
12. Walking without aid		.16	.18	.24	.36	.47	.50	.58	.70	.77	.77			
13. Tap test		.15	.17	.22	.34	.44	.47	.55	.66	.73	78	.77		
14. Step-ups		.13	.15	.19	.29	.38	.40	.47	.56	.62	.71	.81	.86	

N.B Item 1 - sitting with upper limb support - showed zero variance as all the subjects were able to perform it, it was not therefore included in the analysis.

SECTION 4.2.3 RESPONSIVENESS TO CHANGE.

4.2.3.1 Introduction and Method.

Responsiveness to change is an important psychometric property of any outcome measure. The lack of outcome measures with sufficient sensitivity to detect short-term changes has previously been identified as a factor limiting research into the effectiveness for physiotherapy for people with stroke (Partridge 1994) and would be a desirable feature of any new outcome measure. In the past, responsiveness (or sensitivity) to change has usually been assessed by using the measurement tool in a situation where change is expected to occur, most usually between admission and discharge (see Hsieh et al 2000; Malouin et al 1994; McKnight & Rockwood 1995 for example). If a change in score is detected then the measure is said to be sensitive or responsive to change. This method is of limited usefulness if testing whether a measure can detect changes due to a specific therapy intervention. The main difficulty is in deciding the time period over which the change would be expected to occur. If too short a period were chosen then slowly recovering subjects would not show a change and the measure would appear insensitive. If too long a period were chosen it would be difficult to differentiate the changes due to the therapy under investigation from the effects of other aspects of rehabilitation. A change may be detected but one could not say whether it was due to the therapy.

Another method to detect change is to statistically assess responsiveness by calculating the measurement error, which is based on the variability of performance. It estimates how large a change in score is required to demonstrate a 'true change' in performance, rather than variation due to random error (Streiner & Norman 1995; Beckerman et al 1996). This was a more suitable method for this study because it was independent of the subjects' level of ability and the time period over which it was measured.

Three sources of error were investigated: variation within a single testing session (within-session error); between testing sessions (test-retest error); and between testers (inter-tester error). The measurement error between the two sets of measurements was calculated for the overall scale, the sub-sections (sitting, standing and stepping) and the individual performance tests. Nominal data is not suitable for this calculation and so levels 1, 4 and 7 were not included. The measurement error is calculated from the between-subject mean square error (the MSE) of a one-way Anova (Bland 1995) where:

$$\textit{the measurement error} = 1.96x\sqrt{2xMSE}$$

The percentage of the measurement error to the overall mean score (measurement error/ mean score x 100) was next calculated to allow meaningful comparison between the performance tests. Finally, the sources of error were calculated. Systematic error refers to general trends for change in score (either positive or negative) due to practice or fatigue, e.g. Atkinson & Nevill (1998) and is calculated from the mean of the difference (χ_{diff}) between the two measurements (Altman & Bland 1983). Random error is a measure of the variation between repeated tests without trends that are not controlled by the method and include inherent biological or mechanical variability (Atkinson & Nevill 1998). It is calculated from the standard deviation of the differences (sd_{diff}) between the two set of measurements (Altman & Bland 1983). The closer χ_{diff} is to zero and the smaller the sd_{diff}, the less error is found.

Subjects were recruited from participating hospitals using the inclusion criteria and recruitment methods described earlier in Section 4.2.1. The Brunel Balance Assessment was used to assess the subjects' balance as described in Appendix III. The author assessed them wherever they received their physiotherapy treatment: at the hospital bedside; in the physiotherapy treatment area, or in their own home. Data to test the responsiveness to change was collected from thirty-seven people (subjects 1-37). Their demographic details are described in Figure 4.2. This data was also used to test reliability (Section 4.2.2).

In order to test the within-session error, the scale was repeated three times in one session. To test the test-retest error, the scale was repeated once the next day, or before the next physiotherapy session (for out-patients) and the scores for the two days compared. The author collected all the data for these two aspects of responsiveness. She was accompanied on one of the testing sessions by another physiotherapist to test the inter-tester error. Whether the other physiotherapist attended the first test or retest session was at the convenience of the subject and other physiotherapist. The other physiotherapist was one of two of the author's colleagues at Brunel University. One was a full-time lecturer in neurological physiotherapy who had qualified 11 years previously and had eight years of neurological experience. The other was a lecturer/ practitioner in neurological physiotherapy who taught at Brunel University and was a clinical specialist at the Regional Rehabilitation Unit at Northwick Park Hospital. She had qualified nine years previously and had five years specialist neurology experience. To test the inter-tester error the two testers (the author and one other) scored the Brunel Balance Assessment simultaneously. One person instructed the subject and the other observed and noted the subject's scores. The author or the other physiotherapist alternately took the instructor or observer role. Simultaneous scoring was done so that the only source of variation was in the scorers' judgements. If the physiotherapists had instructed subjects on separate occasions, test-retest or within session variation would have contaminated the results.

The measurement error of the performance tests were calculated first and these were then used to calculate a minimum change that could be said to reflect a true change in performance. This score was used as the minimum score to 'pass' the corresponding item on the scale. No advice on how to do this could be found during extensive searching of the literature. So in the absence of any information to the contrary, the author chose the largest measurement error (rounded up) from all three sources of error (within session, test-retest and inter-tester). This would give a single score which was important for use as a pass/fail cut-off point for the scale and a generous estimate of the minimum true change.

4.2.3.2 Results

Measurement Error of the Individual Performance Tests

The measurement error for the performance tests is shown in Table 4.14. The measurement error ranged from zero for the Step-Up Test to 44% for inter-tester error in the Weight Shift test. Test-retesting showed the largest measurement error for five of the ten performance tests. For the Weight-Shift Test, walking without an aid and the Tap Test, the within-session and test-retest measurement error was approximately equal (2% or less, difference between them). For all tests except the Weight Shift test the inter-tester error was the least source of error.

Table 4.14. Measurement Error for the Individual Performance Tests

Test	Mean (sd)	Measurement Error		
		Within-session (% mean)	Test-retest (% mean)	Inter-tester (% mean)
Sitting Arm Raise (n=35)	9.8 lifts (3.5)	2.5 (25%)	2.9 (30%)	1 (10%)
Sitting Fwd Reach (n=34)	26.3 cm (10.6)	5.3 (20%)	10.6 (40%)	2.1 (8%)
Standing Arm Raise (n=28)	9.5 lifts (3.4)	1.2 (16%)	2.4 (25%)	0.6 (6%)
Standing Fwd Reach (n=25)	16.4cm (7.5)	4.2 (25%)	6.6 (40%)	2.5 (15%)
5m Walk with aid (n=21)	14.5s (10.7)	3.6 (25%)	4.3 (30%)	1.67 (11%)
Weight shift (n=18)	5.9 transfers (1.7)	1.3 (22%)	1.2 (20%)	2.6 (44%)
5m Walk without aid (n=14)	2.05s (1.3)	0.67 (33%)	0.63 (31%)	0.36 (18%)
Tap Test (n=13)	7.7 taps (2.2)	1.8 (23%)	1.7 (22%)	0.5 (6%)
Step-ups (n=13)	3.4 step-ups (1.2)	0.94 (28%)	0 (0%)	0 (0%)

The minimum change that could be said to reflect a true change in performance of the individual performance tests (the largest measurement error from all three sources, rounded up) and that would be used as pass/fail criterion for the overall Brunel Balance Assessment are listed below:

- Sitting Arm Raise: 3 lifts — Sitting Forward Reach: 11cm
- Standing Arm Raise: 3 lifts — Standing Forward Reach: 7cm
- Walking with an aid: 4.3s — Weight Shift Test: 3 shifts
- Walking without an aid: 0.7s — Tap Test: 2 taps
- Step-ups: 1 step

The Overall Scale and Sub-Sections

The individual subjects' scores on the Brunel Balance Assessment are shown in Table 4.15. The median score on the overall scale was 10 (Inter Quartile Range 4-8, range 1-12). There was no variation in score for the overall scale, or the sitting, standing or stepping sections in within-session, test-retest or inter-rater testing. This indicated that there was no measurement error and so any change in score would indicate a change in performance

Table 4.15 Individual total scores in testing measurement error of the Brunel Balance Assessment

Subject	Within -session			Test-Retest		Inter-rater	
	S1	S2	S3	Test	Retest	Rater 1	Rater 2
1	12	12	12	12	12	12	12
2	5	5	5	5	5	5	5
3	2	2	2	2	2	2	2
4	4	4	4	4	4	4	4
5	6	6	6	6	6	6	6
6	4	4	4	4	4	4	4
7	8	8	8	8	8	8	8
8	12	12	12	12	12	12	12
9	12	12	12	12	12	12	12
10	12	12	12	12	12	12	12
11	10	10	10	10	10	10	10
12	12	12	12	12	12	12	12
13	3	3	3	3	3	3	3
14	12	12	12	12	12	12	12
15	3	3	3	3	3	3	3
16	12	12	12	12	12	12	12
17	12	12	12	12	12	12	12
18	8	8	8	8	8	8	8
19	6	6	6	6	6	6	6
20	6	6	6	6	6	6	6
21	8	8	8	8	8	8	8
22	3	3	3	3	3	3	3
23	1	1	1	1	1	1	1
24	4	4	4	4	4	4	4
25	12	12	12	12	12	12	12
26	3	3	3	3	3	3	3
27	7	7	7	7	7	7	7
28	10	10	10	10	10	10	10
29	7	7	7	7	7	7	7
30	8	8	8	8	8	8	8
31	4	4	4	4	4	4	4
32	9	9	9	9	9	9	9
33	12	12	12	12	12	12	12
34	12	12	12	12	12	12	12
35	12	12	12	12	12	12	12
36	12	12	12	12	12	12	12
37	12	12	12	12	12	12	12

Sources of Error

The systematic and random error for the performance tests are shown in Table 4.16. In most cases the systematic error was lower than the random error, showing that learning or fatigue had little effect on performance. Exceptions were in the within-test error of the Standing Forward Reach and 5m Walk Test (with or with an aid). Both showed a practice effect - scores improved as subjects became familiar with the test. The forward reach tests (in sitting and standing) also showed quite high degrees of random error. This indicated that some variables were not fully controlled in the standardisation of the testing.

Table 4.16. Sources of error for the performance tests

	Within-Test Reliability		Test-Retest Reliability		Inter-tester Reliability	
	S. E. (x_{diff})	R. E. (sd_{diff})	S. E (x_{diff})	R. E (sd_{diff})	S. E (x_{diff})	R. E (sd_{diff})
Sit. Arm Raise (lifts)	0.4	1.2	0.3	1.4	0.1	0.5
Sit. Fwd Reach (cm)	1.5	2.31	0.27	5.5	0.15	1.1
St. Arm Raise (lifts)	0.1	0.7	0.3	1.2	0.1	0.3
St. Fwd Reach (cm)	1.56	1.50	0.68	3.48	0.20	1.15
Walking with aid (s)	0.5	0.33	0.1	0.25	0.4	0.5
Wt Shift (no. of transfers)	0.2	1.2	0.1	0.8	0.9	0.9
Walking without aid (s)	0.5	0.33	0.1	0.25	0.4	0.5
Tap Test (no. of taps)	0.7	0.9	0.1	0.9	0.1	0.3
Step-Up Test (steps)	0.2	0.4	0.0	0.0	0.1	0.3

S.E = systematic error   R.E.= random error

4.2.3.3 Discussion

Measurement error of the ordinal scales

The ordinal scales of the Brunel Balance Assessment (the overall scale and sitting, standing and stepping sub-sections) showed no measurement error, indicating that any change in score would indicate a true change in performance. This is very low but not out of keeping with other outcome measures which test mobility or balance disability. The Rivermead Mobility Index (RMI) is accurate for changes of two items or more (Collen et al 1991), and the balance section of the Fugl-Meyer Assessment (FMA) is accurate for changes of more than 4 points (Beckerman et al 1996).

To suggest that an outcome measure, even an ordinal scale, has a measurement error of zero is extreme and should be questioned. However when the actual tasks that the testers have to judge is considered it is not surprising that very low measurement error is found. In the Brunel Balance Assessment the actual judgements are simple, very precise and objective. The tester merely has to decide whether the subject has lifted their arm three times, maintained a position for thirty seconds, performed more than one step-up etc. This does not require any particular clinical skill or detailed ability and so it is not surprising that a high degree of agreement about the subjects' ability to perform such tasks was discovered. In contrast, previous outcome measures have included more subjective or less precise measures. The RMI items are all objective but not very precise as they ask the subject "Do you ...?" (Collen et al 1991), while the FMA contains non-standardised, subjective items such as "parachute reaction to the affected side" (Sandford et al 1993).

There are number of limitations to this study of measurement error and these are discussed with the results of reliability testing in section 4.2.3.3.

Measurement error of the performance tests

Goldie et al (1990) used measurement error to assess two interval/ratio tests. They reported an error of twenty-six lifts per minute for the Arm Raise Test in standing, which is greater than the measurement error of three lifts over 15s found in this study. This may be because they measured over a longer period, in which fatigue may have become a factor. While piloting the Brunel Balance Assessment, 15s was chosen as the measurement period as a longer period was tiring. Alternatively, it could be because Goldie et al asked subjects to lift the arm to shoulder height only (rather than to full elevation). This would have the advantage of controlling for people with limited range, but would be harder to determine accurately and harder for subjects to reproduce accurately, which would increase the random error.

Goldie et al (1990) also reported the measurement error of an instrumented version of the Weight Shift test using a limb-load monitor over 60s. This produced an intra-session error of 19 and a test-retest error of 36, which contrasts with the scores found over 15s in this study (2.5, and 1.3 respectively). There are several possible reasons for this. The ratio-level measure from the limb-load monitor would be a more sensitive measure of weight-transfer that would detect a more variable performance. Fatigue may have been a factor over the longer measurement period making the subjects' performance more variable. The task Goldie et al (1990) asked their subjects to perform was also more precise, which may have contributed more variance. They asked subjects to transfer 50% of their weight on to the limb load monitor, whereas subjects in this study were merely asks to transfer their weight from one leg to another. A final difference was that Goldie et al (1990) used a slightly different method to calculate the measurement error (limits of agreement), which may further account for the differences.

Collen et al (1990) reported a test-retest variability of 25% in gait speed (from the difference between tests/ slowest test x 100). Applying the same calculation on the data from this study gives a mean variability of 12%. This difference may be due to the different testing periods. Collen et al had a period of several weeks between test and retest, while in this study it was a matter of days.

For the Brunel Balance Assessment, the test-retest error was the main source of variability. It is not surprising that the test-retesting shows the most variability since measurement on two separate days would be most vulnerable to random error. For several tests the within-session error was similar to the test-retest error and this is probably a reflection of variability in subject performance or the tester's accuracy, since other factors would be unlikely to change in the few seconds between repeating the tests. The inter-rater error was the least (except for the Weight Shift test). As the items of the Brunel assessment are precise and objective there would be little scope for variation. The two testers measured the subject simultaneously. This method was chosen deliberately to reduce the random error that would contaminate the results, if the inter-tester error was assessed by testers assessing the subject on two different occasions, although it may have inflated the scores. There was no obvious reason for the greater variability shown for the Weight Shift Test.

Sources of Error

In most of the performance tests, the random error was higher than the systematic error. A relatively high random error is not surprising in people with impaired motor control, as inconsistency of performance is a feature of unskilled movement (Shumway-Cooke & Woolacott 2000). Random error has also been reported as a main source of error in platform tests (Corriveau et al 2000, Geurts et al 1993, Dickstein & Dvir 1993, Liston & Brouwer1996) and walking speed (Evans et al 1997). Little systematic error was identified in the performance tests except in the within-test error of the Standing Forward Reach and the 5m Walk Test. This indicates that for most of the tests the effects of learning/ practice or fatigue were negligible. For the Forward Reach and 5m walk tests. The subjects' performance tended to improve during the testing session as they became familiar with the test. Recommended ways to overcome this problem are to repeat the test several times and take the best score, the last score or the average score (Evans et al 1997, Goldie et al 1999). To decide which method to use for further tests of the ratio measures (sitting forward reach, standing forward reach and 5m walk tests) the measurement error for all the various combinations was compared by comparison with the score for a single test using the trials for the within-session testing. The possible combinations were to take: the second or third score; the best of the first and second, or all three scores; or the average scores of the first and second, first and third, second and third or all three scores. The scores are detailed in Table 4.17. For all three ratio tests taking the average of the first and second score produced the least measurement error and so this was used for subsequent testing of the Brunel Balance Assessment.

Table 4.17. Measurement error for different scoring combinations for the ratio tests.

Combination	Measurement error		
	5m walk test (s)	**Standing fwd reach (cm)**	**Sitting fwd reach (cm)**
Test 2	7.54s	10.18cm	9.41 cm
Test 3	7.99s	15.28cm	17.14 cm
Best of test 1 or 2	8.54s	11.36 cm	13.78 cm
Best of test 1, 2 or 3	14.27s	16.86 cm	22.52 cm
Mean of test 1& 2	*3.76s*	*5.10 cm*	*4.72 cm*
Mean of test 1 & 3	4.00s	7.64 cm	8.57 cm
Mean of test 2 & 3	7.79s	12.74 cm	13.28 cm
Mean of tests 1,2, & 3	8.1s	8.49 cm	8.85 cm

In all cases the combination is compared with the score for one single test.
Italics = combination with the least measurement error

Few other studies have reported sources of error in performance tests. Goldie et al (1990) reported a learning effect in a test battery that included the standing Arm Raise Lift test and instrumented Weight Shift test. Each of the tests was described as novel to the patient, which may explain the learning effect. This effect was not found in this study, but the tasks were deliberately selected to be familiar to the subject, either from everyday life or from tasks they would practice in therapy.

Evans et al (1997) specifically examined the sources of error while measuring gait. They used footswitches to assess gait over 10m in a group of subjects who were more able walkers than in this study and used a different method to calculate error (standard error of the difference). Differences between the two studies were inconsistent; the random error within tests was similar, systematic error within tests was less in the Evans et al study, but there was greater random error between test-retesting. It is possible that these differences are due to the inherent variability of gait patterns rather than any consistent differences between the studies.

In the present study, the Forward Reach tests showed a learning effect and quite high random error, particularly between test-retesting. Duncan et al (1990) reported a coefficient of variation (sd of 3 trials/mean of 3 trials x 100) of 2.5% for the Standing Forward Reach when tested with healthy adult volunteers (people with stroke were excluded). Using this method for the within session data produces a coefficient of variation of 6.1%. The most obvious explanation for the difference in score is a greater variability of performance in the stroke patients of this study in comparison with the healthy adults of the Duncan et al study. The relatively high degree of error highlights the need to familiarise the subject with the test (to reduce systematic error) and to carefully standardise the procedure, particularly the position of the subject and measurement ruler (to minimise the random error).

The ratio measures showed higher error overall than the interval measures. The only tests to show a learning effect were ratio measures and they tended to show more random error. This is probably a reflection of their greater sensitivity, as the increment between units is smaller.

4.2.3.4 Conclusion

The responsiveness of the Brunel Balance Assessment to change has been assessed by calculating the measurement error. Where comparison is possible, the Brunel assessment showed similar or less variability than other similar outcome measures. The ordinal Brunel scales have been found to be sensitive, such that any change in score could be attributed to a 'true' change in performance. For the individual performance tests, scores have been calculated which account for different sources of error and represent the change required to represent a 'true' change in performance. This could be used for research, or clinically, to assess the effects of treatment over any period.

Sources of error for the performance tests have also been identified. The interval tests showed less error (random or systematic) than the ratio tests, as would be expected due to the larger differences between increments for the interval tests. Several ways of scoring the ratio tests were examined to identify which produced the least measurement error. The most effective method was to repeat the test and take an average score of the two attempts. This method was adopted for subsequent testing of the ratio measures (Forward Reach tests and 5 m Walk tests). The protocol of the Brunel Balance Assessment was adjusted accordingly.

Two aims in developing the Brunel Balance Assessment were to develop a measure that was sensitive to changes in performance and was suitable for a wide range of severity in performance. These aims have been fulfilled. By using measurement error, rather than 'change over time' to assess sensitivity, a change in performance above the minimum score can be said to represent a true change regardless of the time span over which it occurred, the subjects' abilities, or the rate of recovery.

There are few examples of the use of measurement error in the rehabilitation literature yet it simple to calculate and, as the values are in the same units as the original measurements, it gives useful and meaningful information for clinical practice as well as research. Its wider use is recommended.

SECTION 4.2.4 RELIABILITY

Having established scalability and responsiveness to change, the next stage in developing the psychometric properties of the Brunel Balance Assessment was to assess reliability. The scale and the individual performance tests were tested. Three aspects of reliability were tested:

- Inter-rater reliability - reproducibility of score when tested by two assessors
- Test-retest reliability - reproducibility of score when repeated on two separate occasions
- The within-session or intra-session reliability - reproducibility of score when the test is repeated several times on the same occasion.

4.2.3.1 Method

Data from 37 subjects was collected as described in section 4.2.1 and Figure 4.2a&b and the Brunel Balance Assessment was used to assess the subjects' balance as described in the instruction manual (Appendix III). The subjects were the same as those presented in section 4.2.3 (Responsiveness to change) and the data for these two sections were collected at the same time. They are presented separately merely for clarity and ease of presentation.

Data to test the reliability of the scale and the individual tests were collected simultaneously. Each performance test also formed an item or level on the scale as explained in section 4.1.1. For each performance test there is a minimum score to 'pass' that item of the scale based on the measurement error (see results section of measurement error of performance tests, Page 141-2). The subjects performed each test in turn and their score was noted. If they reached the minimum score, the subject passed that level of the scale and testing progressed to the next level. If the subject could not perform the test or reach the minimal score, two more attempts (three in total) would be made. If after three attempts a 'pass' had not been achieved then testing ceased. Further items did not need to be tested because the scale forms a Guttman-type scale.

To assess inter-rater reliability the results from two testers were compared. The two testers assessed subjects simultaneously to reduce the effects of random error from repeating the test on two different occasions. The testers alternated between instructing the subject, and observing. The testers were the author (ST) and another person. The other being one of two colleagues at Brunel University, they are described in Section 4.2.3.1 (responsiveness to change section).

Test-retest reliability was assessed by testing the subjects' balance with the Brunel Balance Assessment on two separate occasions. As most of the subjects were in the acute stages post-stroke it was important to test and retest quickly so that changes due to spontaneous recovery or treatment did not interfere with the results. Wherever possible, retesting took place the next day before the subject had another physiotherapy session. This did happen in most cases, but was not always possible. When it was not possible, retesting was arranged at the subject's earliest convenience. Within-session reliability was assessed by repeating each test three times during one session, then the results for the first and third test were compared. The first and third attempts were compared as the first and last test was expected to show the largest change in score.

Analysis

The Kappa Co-efficient, and weighted Kappa was used for nominal and ordinal data respectively (Cohen 1968; Cohen 1960). These co-efficients calculate the level of agreement between sets of data in proportion to the level of agreement expected by chance alone, where a score of 1 indicates perfect agreement. The weighted Kappa also accounts for the size of the disagreement between the observations.

Intra-class correlations (ICC) were used to assess the reliability of the interval and ratio data (Shrout & Fleis 1979). The ICC is an index obtained by taking the variance between and within subjects into account. It gives a single figure that reflects the consistency and agreement between sets of data. The closer the coefficient is to 1, the higher the reliability. There are six different formulae for the ICC's. In this case the (3,1) formula was used for the inter-rater reliability, the (1,1) formula was used for the test-retest reliability and the (1,3) formula was used for the within-test reliability (Rankin & Stokes 1998).

4.2.3.2 Results

The Ordinal Scales

The individual subjects' scores on the ordinal scale are the same as shown in section 4.2.2.2 (Responsiveness to change, Table 4.14). The median score on the overall scale was 10 (IQR 4,8 Range 1-12). The overall scale and the sitting, standing, and stepping sections showed a very high degree of reliability - there was 100% agreement on the test-retest, within-test, and inter-tester reliability (k=1).

The Performance Tests

Nominal Data

All the nominal data - level 1 (sitting with upper limb support), level 4 (standing with upper limb support), and level 7 (static step standing without support) showed 100% agreement on the pass/fail between testers, test-retest and within-test reliability giving, in all cases, a kappa co-efficient of 1. All subjects passed level 1, thirty-two (82%) passed level 4, and twenty-three (59%) passed level 7.

The interval and ratio data.

The mean scores for each of the performance tests are shown in Table 4.17. The results of the reliability testing for each of the interval and ratio tests are shown in the Table 4.18. High levels of reliability were found with the intra-class correlations ranged from 0.88 (for test-retest of the Weight-Shift test) to 1 (for the test-retest and inter-tester reliability of the Step-Up Test). Generally the test-retest reliability was lowest and the inter-tester reliability the highest.

Table 4.18. Scores in the interval and ratio performance tests, showing the mean,(standard deviation) and range of scores.

Performance Test	No.	Within Session reliability			Test -retest		Inter-rater	
		S1	**S2**	**S3**	**Test**	**Retest**	**Rater1**	**Rater 2**
Sititng AR	35	9.5 (3.0) 5-16	9.8 (3.5) 4-17	9.9 (3.5) 4-18	9.8 (3.4) 5-18	10.1(4) 3-18	9.7 (3.4) 5-18	9.9 (3.4) 5-18
Sitting FR	34	25.5 (10.2) 7-45	26.3 (10.7) 7-49	27.0(10.5, 7-50)	26.8(10.3) 7-44	26.5 (10.8) 10-45	26.6 (10.9) 12-44	26.7 (11.0) 11-45
Standing AR	28	9.4 (3.3) 2-17	9.5 (3.5) 3-17	9.5 (3.4) 3-17	9.5 (3.4) 3-17	9.8 (3.6) 5-17	9.4 (3.4) 3-17	9.3 (3.3) 3-17
Standing FR	25	15.76 (7.08) 5-30	16.80 (7.38) 5- 33	17.32 (7.54) 5- 33	16.68 (7.48) 5-33	16.0 (7.73) 6- 28	16.20 (7.71) 6-33	16.0 (7.82) 6-34
Walk with aid	21	15.03 (10.46) 5.35-50.53	14.19 (9.57) 5.28-46.84	14.14 (10.52) 5.34- 50.56	14.15 (10.16) 5.37- 50.56	14.39 (11.39) 4.87- 55.71	14.38 (11.27) 5.37-55.71	14.19 (11.3) 5.72-55.88
Weight-shift	17	5.4 (1.8) 3- 10	5.9 (2.1) 2-11	6.3 (2.0) 4-12	6.1 (1.8) 4-11	5.9 (1.7) 3-10	5.9 (1.4) 4-10	5.8 (1.3) 4-8
Walk without aid	14	2.19 (1.48) 0.92-5.27	2.09 (1.4) 0.93-5.30	2.05 (1.45) 0.79-5.42	2.02 (1.25) 0.96-4.58	1.96 (1.28) 0.97-4.9	2.06 (1.37) 0.87-4.9	2.04 (1.4) 0.92-5.27
Tap Test	13	7.1 (1.9) 4-11	7.6 (2) 5-11	7.9 (2.4) 5-13	7.8 (2.1) 5-11	7.9 (2.3) 5-12	7.9 (2.4) 5-13	7.8 (2.3) 5-13
Step-ups	12	3.3 (1.3) 2-5	3.2 (1.0) 2-5	3.5 (1.3) 2-6	3.5 (1.3) 2-6	3.5 (1.3) 2-6	3.5 (1.3) 2-6	3.5 (1.3) 2-6

AR = Arm Raise test, FR = Forward Reach Test

Table 4.19. The Intra-class correlations (with 95% confidence intervals) for the reliability of the performance tests.

	Within-Test Reliability	Test-Retest Reliability	Inter-tester Reliability
Sitting Arm Raise	0.96 (0.92, 0.98)	0.96 (0.92, 0.98)	0.99 (0.99, 0.99)
Sitting Fwd Reach	0.98 (0.97, 0.99)	0.93 (0.86, 0.96)	0.99 (0.99, 1.0)
Standing Arm Raise	0.98 (0.97, 0.99)	0.93 (0.86, 0.96)	0.99 (0.99, 1.0)
Standing Fwd Reach	0.98 (0.95, 0.99)	0.95 (0.88, 0.98)	0.99 (0.99, 1.0)
Walking with an aid	0.99 (0.98, 1.0)	0.99 (0.97, 1.0)	1.0 (0.99, 1.0)
Weight Shift	0.86 (0.60, 0.95)	0.91 (0.77, 0.97)	0.88 (0.68, 0.96)
Walking without an aid	0.99 (0.96, 1.0)	0.99 (0.95, 0.99)	1.0 (0.99, 1.0)
Tap Test	0.93 (0.78, 0.98)	0.96 (0.88, 0.99)	1.0 (0.99, 1.0)
Step-Up Test	0.97 (0.89, 0.99)	1.00 (N/A)	0.97 (0.89, 0.99)

N/a = not applicable. There was no change in any subjects on test-retesting.

4.2.3.3 Discussion

Can a measure be 100% reliable?

The test-retest, inter-tester and within-test reliability of the overall scale, the sitting, standing and stepping sections and the performance tests was found to be high. In fact, the results suggest 100% reliability in score and the results of Section 4.2.2 suggest a measurement error of zero. These are unusually high and should be considered critically.

As discussed in Section 4.2.2.3 (discussion of responsiveness to change) the actual task that is judged to pass or fail a subject on each level of the scale is simple, very precise and objective. The tester merely has to decide whether the subject has lifted their arm three times, maintained a position for thirty seconds, performed more than one step-up etc. This does not require any particular clinical skill or detailed ability and so it is not surprising that a high degree of agreement about the subjects' ability to perform such tasks was discovered. Previous outcome measures have included more subjective or less precise measures. The Rivermead Mobility Index items are all objective but not very precise as they ask the subject "Do you ...?" (Collen et al 1991), while the Fugl- Meyer Assessment contains non-standardised, subjective items such as "parachute reaction to the affected side" (Sandford et al 1993) and the MAS requires subjective judgements such as "sit well forwards with weight evenly distributed" (Carr et al 1985). This may account for the slightly higher scores with the Brunel scale than some ordinal scales. The Berg Balance scale also uses precise objective criteria and reports very high reliability (Berg et al 1995).

Limitations of the Study

There are a number of limitations that need to be borne in mind when considering the results of this study. Firstly the scope of testing was limited. There were thirty-seven subjects and inevitably relatively few had impaired sitting balance. This means that most of the subjects in the sitting section passed all the levels. It is assumed that similar levels of accuracy would be found in more impaired subjects. This is not an unreasonable observation because the five subjects with impaired sitting balance had perfect agreement scores on the ordinal and nominal tests and the score achieved in this assessment is similar or better than that found in other assessments. Similarly, relatively few subjects were able to perform the hardest tests. There is no recognised method to calculate the number of subjects needed for studies testing reliability, but the numbers included, and the levels of agreement, in the present study were similar to other published work (Goldie et al 1990, Collen et al 1990).

The inter-tester reliability in particular was limited in scope. The agreement between two experienced physiotherapists, when testing the subject simultaneously, was assessed. Simultaneous testing was used to reduce the random error involved in testing on different occasions, so that any error could be attributed to differences between testers. This does not reflect the clinical setting however, when separate testing would take place. The involvement of only two experienced physiotherapists is also insufficient to make generalisations. Further testing with physiotherapists with different degrees of experience and other health care professionals is needed before the results can be generalised. This was not possible in the present study due to limited resources available.

Taking the test-retest measurements close together was necessary to prevent changes in results due to recovery or therapy in acute stroke patients. However the tester or subject may have influenced the results if s/he could remember the scores from the previous testing occasion. An attempt was made to prevent bias during re-testing. For instance, the scores for the first testing were not available to the tester until re-testing was complete, but sub-conscious bias is a possibility that would have artificially inflated the levels of agreement.

The final limitation is in the selection of subjects. As discussed above there were relatively few subjects at the extreme ends of ability. Although the numbers are similar to other studies (Collen et al 1990; Goldie et al 1990,), further testing with larger numbers would allow greater confidence in the findings. An aim in recruiting the subjects was that the study population should represent people with stroke receiving out-patient physiotherapy. However, some patients were excluded by the need to obtain informed consent or to include only people with balance problems due to their stroke. The Brunel Balance Assessment may be a less satisfactory measure of balance disability with these sub-groups.

4.2.3.4 Conclusion

The results of this study show that the Brunel Balance Assessment has a high level of reliability when used by experienced physiotherapists with people who have balance disability caused by a stroke. This includes people with a wide range of severity and duration of hemiplegia.

SECTION 4.2.5 VALIDITY

4.2.5.1 Method

The final aspect of scale development to be assessed was the validity. Three aspects of validity were tested: *concurrent validity* in which scores for the Brunel Balance Assessment were compared with established tests of the same construct; *criterion related validity* in which Brunel Balance Assessment scores were compared with established tests of related constructs; and *discriminant validity*, in which the sitting, standing or stepping sections were used to differentiate levels of ability.

Scales identified as the 'best of the bunch' in Chapter 2 were used for comparison with the Brunel Balance Assessment to test the concurrent validity. These were:

- the sitting section of the Motor Assessment Scale (MAS, Carr et al 1985)
- the Rivermead Mobility Index (RMI, Collen et al 1991)
- the Berg Balance Test (Berg et al 1992)

Criterion related validity was assessed by comparing the Brunel Balance Assessment score with:

- the leg score of the Motricity Index (MI, as a measure of motor impairment, Demeurisse et al 1980)
- the Rivermead Mobility Index (RMI, as a measure of motor disability Collen et al 1991)
- the mobility section of the Nottingham Extended ADL score (EADL, as a measure of motor handicap, Gladman et al 1993). This was only used with subjects who were living at home as in-patients would not have experience of their levels of community-based handicap since their stroke.
- the Barthel Index (as a measure of disability in the activities of daily living, Collin et al 1988).

All the comparative tests listed above were subsequently referred to collectively as the 'other tests'.

Discriminant validity was tested by dividing the subjects according to their level of balance ability and comparing their scores on the 'other tests'. Balance ability was divided into three groups. There were people who could sit but not stand (scoring 1-3 on the Brunel scale, referred to subsequently as 'sitters'); people who could stand but not step (scoring 4-6 on the Brunel scale, referred to subsequently as 'standers'); and people who could sit and stand but had impaired stepping or walking (scoring 7+ on the Brunel scale, referred to subsequently as 'steppers'). Details of the 'other tests' and the way in which they were performed are detailed in Appendix IV.

Analysis

Spearman's rho correlations were used to compare the Brunel Balance Assessment and the performance tests, with the other tests for the concurrent and criterion related validity and a one-way ANOVA was used to assess discriminant validity.

4.2.5.2 Results

Subjects

Data from 55 subjects was collected as described in Section 4.2.2 and Figure 4.2 and the Brunel Balance Assessment was used to assess the subjects' balance as described in the instruction manual (Appendix III). The data from these subjects was also used for the mathematical modelling in Chapter 5 and so more detail of their abilities and impairments was collected than the other subjects in this scale development study. This is detailed below in Table 4.20. There were 32 men (58%) and 23 women, 29 (53%) had a left hemiplegia and 26 had a right hemiplegia

Subjects were also grouped according to their balance ability. The demographic details of people in the different balance groups are given below (Table 4.21). Further detail of their performance on the 'other tests' are found in Table 4.23

Table 4.20 Demographic details of the subjects included in the validity testing.

	Mean	Standard deviation	Range
Age (years)	68	12.8yrs	40-88
Time since stroke (weeks)	9 (median) (4,24 IQR)		
Brunel	6.3	3.8	1-12
Berg	24	20.6	0-56
Rivermead	6.4	4.5	1-14
Motricity Index (leg)	47.9	29.5	0-92
Barthel Index	11.4	5.7	1-20

Brunel = Brunel Balance Assessment, Berg = Berg Balance Test, Rivermead = Rivermead Mobility Index, IQR = inter-quartile range

Table 4.21. Demographic details of the sub-groups

	Sitters (n=17)	Standers (n=18)	Steppers (n=20)
Age (years)	72.9 (sd=12.7)	65.1 (sd=13.6)	67.25 (12.9)
Sex	7M / 10F	11M / 7F	13M / 7 F
Side of hemiplegia	9L/ 8R	11L / 7R	9L / 11R
Median time since stroke (weeks)	8 (IQR2, 21)	9 (IQR 5.5, 40.5)	8.5 (IQR 4, 14)

M=male, F=female, L=left, R=right, IQR + inter-quartile range, Brunel = Brunel Balance Assessment, Berg = Berg Balance Test, Rivermead = Rivermead Mobility Index, MAS = Motor Assessment Scale, EADL = Extended Activities of Daily Scale

Concurrent and Criterion related validity.

The Brunel Balance Assessment showed high, statistically significant correlations with all the 'other tests' (Table 4.22). This indicated that the Brunel Balance Assessment demonstrated good concurrent and criterion related validity as a test of balance. The sub-scales showed significant relationships with the other tests except the Motricity Index. This indicated that the sub-scales demonstrated good concurrent and criterion related validity as tests of sitting, standing and stepping balance respectively

The correlation between the individual performance tests of the Brunel Balance Assessment and the other tests are shown in Table 4.23. In this Table, the data for people who walked with (n=8) and without an aid (n=13) has been combined and is presented as '5m walk' (n=21). This was done to diminish the effect of the relatively small numbers who completed each test. Significant correlations between the items and other measures were found in most cases, indicating that the individual tests were valid tests of balance disability. The weight shift test showed less significant correlations than the other tests and correlations with the Motricity Index and Berg scale did not reach significance.

Table 4.22. Spearman's correlation co-efficients between scores for the overall Brunel Balance Assessment and its sub-sections and the 'other' tests (n=55).

	Correlation with the Brunel Balance Assessment			
	Overall Scale	**Sitting Scale**	**Standing Scale**	**Stepping Scale**
Concurrent Validity				
Berg	0.97**	0.62**	0.86**	0.63**
RMI	0.95**	0.68**	0.68**	0.58*
MAS	0.83*	0.87**	n/a	n/a
Criterion Related Validity				
MI (leg)	0.83**	0.23	0.32	0.39
RMI	0.95**	0.68**	0.68**	0.63*
BI	0.95**	0.71**	0.65**	0.70**
EADL	0.56*	n/a	n/a	0.41

Berg = Berg Balance Test, RMI = Rivermead Mobility Index, MAS = Motor Assessment Scale, MI = Motricity Index, BI = Barthel Index, EADL = Extended Activities of Daily Scale

*N.B **= significant at p<0.01, *= significant at p<0.05. n/a= not appropriate*

Table 4.23. Spearman's correlation between the performance tests and the 'other tests' (n=55).

	Sit Arm Raise	**Sit Fwd Reach**	**Stand Arm Raise**	**Stand Fwd Reach**	**Wt Shift**	**5m Walk**	**Tap Test**	**Step-ups**
MAS	0.33*	0.54**	n/a	n/a	n/a	n/a	n/a	n/a
Berg	0.54**	0.54**	0.36*	0.7**	0.26	-0.64**	0.74**	0.19
RMI	0.53**	0.61**	0.32*	0.57**	0.52*	-0.54*	0.46*	0.66*
MI	0.43**	0.42*	0.34*	0.63**	0.37	-0.64**	0.4*	0.72**
BI	0.41**	0.45*	0.27	0.53**	0.42*	-0.53*	0.74**	0.75**
EADL	N/a	n/a	n/a	n/a	0.54*	-0.35	0.19	0.23
No.	48	46	32	28	18	25	15	12

MAS = Motor Assessment Scale, Berg = Berg Balance Test, RMI = Rivermead Mobility Index, MI= Motricity Index, BI = Barthel Index, EADL = Extended Activities of Daily Scale

*N.B **= significant at p<0.01, *= significant at p<0.05. n/a = not appropriate*
5m walk = walking with aid, and without aid combined.

Discriminant Validity

The Brunel Balance Assessment and its sub-scales

Discriminant validity of the Brunel Balance Assessment was calculated by dividing subjects into the sub-groups; 'sitters', 'standers' and 'steppers' according to their score on the Brunel Balance Assessment. Their details have been described previously in Table 4.21. The scores for these sub-groups in the 'other tests' were compared (Table 4.24). Results from the one-way ANOVA showed significant differences (p<0.001) between the groups in all outcome measures indicating that the Brunel Balance Assessment did have discriminant validity. The separation between the three groups of patients was marked, indicating that the Brunel Balance Assessment may be useful measure to stratify subjects for intervention trials in the future. The full comparison matrix is shown in Table 4.25.

Table 4.24. Mean, standard deviation and range of scores in the 'other tests', for the overall Brunel Balance Assessment and the sub-scales.

Scale	Overall Scale n=55		Sitters (score 1-3) n=17		Standers (score 4-6) n=18		Steppers (score 7-12) n=20	
Brunel	6.4 (4)	0-12	2.1 (1)	1-3	5.2 (8.6)	4-6	11.25 (1)	7-12
Berg	24.7 (21.1)	0-56	2.4 (1.7)	0-5	17.9 (9)	6-34	49.7 (6.1)	30-56
RMI	6.7 (4.7)	0-15	2.1 (1.4)	0-3	5 (1.2)	3-6	12.1 (2.8)	5-15
MAS	5.1 (1.7)	0-6	n/a		n/a		n/a	
MI	50.6 (29.8)	0-100	22.8 (20.8)	0-58	39.6 (7.9)	25-55	84.2 (10.4)	65-100
BI	11.5 (5.9)	1-20	5.2 (2.9)	1-9	10 (1.9)	7-14	18.2 (2.3)	13-20
EADL	11.6 (2.9)	5-15	n/a		n/a		n/a	

Brunel = Brunel Balance Assessment, Berg = Berg Balance Test, RMI = Rivermead Mobility Index, MAS = Motor Assessment Scale, MI= Motricity Index, BI = Barthel Index, EADL = Extended Activities of Daily Scale

Table 4.25. Comparison of performance of sitters, standers and steppers on the 'other tests' using one-way ANOVA with Bonnferoni post-hoc test.

Dependent Variable	Balance level		Mean difference	Std error	Signif	95% CI Lower	Upper
Berg	Sitting	Standing	15.59	2.17	<-0.001	20.95	10.23
		Stepping	47.35	2.11	<-0.001	52.58	42.12
	Standing	Stepping	31.76	2.08	<-0.001	36.91	26.61
RMI	Sitting	Standing	2.89	0.67	<-0.001	4.55	1.22
		Stepping	9.98	0.66	<-0.001	11.61	8.36
	Standing	Stepping	7.1	0.65	<-0.001	8.7	5.5
Barthel	Sitting	Standing	4.82	0.81	<-0.001	6.82	2.82
		Stepping	12.96	0.79	<-0.001	14.92	11.01
	Standing	Stepping	8.14	0.78	<-0.001	10.07	6.22
Motricity	Sitting	Standing	16.85	4.7	<0.02	28.49	5.21
		Stepping	61.39	4.59	<-0.001	72.74	50.03
	Standing	Stepping	44.54	4.52	<-0.001	55.72	33.36

Berg = Berg Balance Test, RMI = Rivermead Mobility Index, Barthel = Barthel Index, Motricity = Motricity Index

The performance tests

The discriminant validity of the performance tests was also assessed. This involved sub-dividing the subjects according to their performance in each of the performance tests and comparing their scores on the various aspects of the 'other tests'. For clarity these are presented as the sitting, standing and then stepping tests.

Sitting Tests

There were two sitting tests - the arm raise and forward reach tests. The discriminant validity of these tests was tested by comparing the scores for each balance ability group (sitters, standers and steppers Table 4.26).

Table 4.26. Mean, standard deviation and range of scores for the sitting tests

	Mean, standard deviation and range			
	Overall Scale	**Sitters**	**Standers**	**Steppers**
Arm Raise	9.1 (2.9) 3-15	6.4 (2.3) 3-9	9.2 (2.4) 5-13	10.4 (2.8) 5-15
Fwd Reach	34.3 (9.2) 15-54	24.8 (7.8) 15-35	33.3 (8.2) 18-47	39.0 (7.3) 8-45

There were significant differences between the groups (p<0.001). In each case, this was between sitters, and standers and/or steppers, but not between standers and steppers (Table 4.27). This indicated that the sitting tests could distinguish between people with sitting balance only, and people who could stand and step.

Table 4.27. Comparison of the Sitting Arm Raise and Forward Reach tests between sitters, standers, and steppers using a one-way ANOVA and Bonneferoni post-hoc test

Dependent Variable	Balance level		Mean difference	Std error	Signif	95% CI	
						Lower	Upper
Arm Raise	Sitting	Standing	2.77	1.0	<0.02	5.25	0.29
		Stepping	3.95	1.0	<0.001	6.38	1.52
	Standing	Stepping	1.18	0.82	<0.48	3.22	0.86
Forward Reach	Sitting	Standing	8.58	3.3	<0.04	16.81	0.36
		Stepping	14.2	3.25	<0.001	22.29	6.1
	Standing	Stepping	5.67	2.53	<0.09	11.9	0.67

A second set of calculations was performed to compare people with static and dynamic sitting balance to see if the performance tests (the Sitting Arm Raise and Forward Reach tests) could discriminate between them. This analysis was confined to people with impaired sitting or standing balance (Brunel Balance Assessment score of 6 or less, n=35). Static sitting balance was defined by a MAS score of three or less (n=11), and dynamic sitting balance was defined by a MAS score of 4 or above (n=24) (See Appendix IV). The details of this subject group are shown in Table 4.28 and the scores on the Arm Raise and Forward Reach Tests are shown in Table 4.29, which show significant differences between people with static and dynamic sitting balance.

These results indicate that the sitting arm raise and forward reach test has discriminant validity and can discriminate between people with impaired sitting balance and people who can stand and step. The tests can also distinguish between people with static and dynamic sitting balance.

Table 4.28. Subjects with impaired standing or sitting balance, showing mean (standard deviation) and range.

Scale	Sitters and standers (BBA <7, n=35)	People with static sitting balance only (MAS<4, n=11)	People with dynamic sitting balance (MAS 4+, n=24)
Sex	18 M / 17W	6 M/ 5 W	8 M/ 6 W
Side	20 R 15L	7 L/ 4 R	9 L/ 5 R
Median time since stroke	9 (IQR 3,30)	8 (IQR 3, 12)	8.5 (IQR 2.75, 20..5)
Age	68.9 (12.2) 40-87	73.8 (12.2) 54-87)	72.5 (13.1) 54-87
Brunel	3.8 (1.8) 1-6	2.4 (1.6) 1-5	4.6 (1.3) 3-6
Berg	10.9 (10.5) 0-34	3.9 (4.8) 0-15	14.7 (10.9) 1-34
RMI	3.7 (1.9) 0-6	2.3 (1.9) 0-5	4.5 (1.4) 3-6
MAS	4.5 (1.9) 1-6	2.0 (0.6) 1-3	5.8 (0.5) 4-6
MI	31.5 (21) 0-84	21.5 (19.4), 0-45	37.1 (20.2) 0-84
BI	8.1 (3.5) 1-14	4.9 (3.2) 1-10	9.8 (2.2) 6-14

M = men, W = women, L = left, R = right.
Brunel = Brunel Balance Assessment, Berg = Berg Balance Test, RMI = Rivermead Mobility Index, MAS = Motor Assessment Scale, MI= Motricity Index, BI = Barthel Index

Table 4.29 Scores in the Sitting Arm Raise and Forward Reach Tests for people with static or dynamic sitting balance.

	People with static sitting balance	People with dynamic sitting balance	P value (95%CI)
Arm Raise (lifts)	5.5 (sd=1.9)	8.6 (sd=2.5)	<0.007 (-28.81, -5.19)
Forward Reach (cm)	20.75(sd=6.5),	32 (sd=7.9).	<0.027 (-5.87, -0.38).

The Standing Tests

There were two standing tests - the arm raise and forward reach tests. The discriminant validity of these tests was tested by comparing the scores for the steppers and standers. The two groups of subjects are described in Table 4.21 and 4.24. There were significant differences between standers and steppers in both tests (Table 4.30), indicating that the standing performance tests had discriminant validity and distinguished between people who could or could not step and walk.

Table 4.30. Mean, standard deviation and range of scores for the standing tests, and comparison between standers and steppers.

	Standers	Steppers	P value (95%CI)
Arm Raise	7.0 (2.2) 3-10	8.8 (2.1) 6-13	0.03 (-3.38, -0.22)
Fwd Reach	17.7 (5.1) 10-24	26.1 (5) 15-36	0.01 (-12.6, -4.0).

<u>**Stepping Tests**</u>

There were four stepping performance tests; the Weight Shift test, combined 5m-walk test, Tap Test and Step-Up test. Discriminant validity was assessed by comparing scores on the stepping tests in people with different levels of functional mobility based their scores on the Rivermead Mobility Index. The mobility functions tested were the ability to: get up and down stairs (RMI score of 8); walk outdoors on uneven ground (with an aid if necessary, RMI score of 12); and get up and down four steps without a rail (RMI score of 14). The mean (sd) scores for people who were able and unable to perform these functions and the p values (95% confidence intervals) of the comparison between those able and unable are shown in Table 4.31.

The 5m Timed Walk Test distinguished between people who could walk outside on uneven ground, and those who could get up and down stairs, while the Tap test distinguished these functions plus walking without an aid. The Step-Up test distinguished people who could get up and down steps from people who could not. The Weight Shift test distinguished between people who could and could not walk outside on uneven ground.

Table 4.31. Scores in the stepping tasks comparing people able and unable to perform functional mobility tasks.

Functional Task		Stepping Tests			
		Weight Shift	5m Walk	Tap Test	Step-ups
Stairs	Able	5.8 (1.8)	*6.3 (2.5)*	*8.9 (1.5)*	4 (1.28)
	Unable	4 (0)	*14.5 (8.3)*	*5.7 (1.7)*	4 (1.28)
	P value	0.18	*p<0.006*	*p<0.004*	1
	(95%CI)		*(-13.68, -2.66)*	*(1.2, 5.2)*	3.2, 4.8
Walking outdoors	Able	*6.25 (1.6)*	*6.34 (2.87)*	*8.86 (1.7)*	4.3 (1.2)
	Unable	*4.3 (1.4)*	*12.46 (8.0)*	*6.5 (2.2)*	3 (1)
	P value	*0.02*	*p<0.03*	*p<0.04*	0.12
	(95%CI)	*4.7, 6.5*	*(-11.5, 0.72)*	*(0.18, 4.53)*	3.2, 4.8
Steps	Able	6.3 (1.5)	7.45 (8.4)	8.2 (2.2)	*5.3 (0.96)*
	Unable	4.8 (1.8)	(4.4)	6.4 (2.1)	*3.4 (0.92)*
	P value	0.06	0.143	0.153	*p<0.008*
	(95%CI)	4.7, 6.5	6.2, 12.5	6.3, 8.9	*0.61, 3.14*

italics = significant at 5%

4.2.5.3 Discussion

Concurrent/ criterion related validity

The Brunel Balance Assessment, sub-scales and individual performance tests demonstrated good concurrent and criterion related validity as a measure of balance disability by significant correlations with most other measures (Table 4.21 and 4.22). There were some exceptions. The Motricity Index did not show significant correlations with the sub-scales. The Extended Activities of Daily Living scale did not show significant correlations with the stepping sub-scale and most of the performance tests. The Motricity Index and the Weight Shift test did not have a significant correlation, nor the Berg scale with the weight shift or step-up test.

A possible reason for the poor correlations could be that balance performance is not the factor limiting performance on the 'other test'. For instance, sensory or perceptual impairment rather than balance disability may impact on motor impairment (measured by the Motricity Index). Other studies have also found poor correlations between measures of balance disability and impairment (Fishman et al 1997; Niam et al 1999), although correlations between lower limb strength and walking are strong (Bohannon 1989; Jorgenson et al 1995).

Other possible explanations are 'mathematical'. Relatively small numbers of people were living in the community and were eligible to complete the Extended Activities Of Daily Living scale because in-patients had been excluded from completing the Extended Activities Of Daily Living Scale as they would not have experience of community living post-stroke (n=14). This would have reduced the chance of strong correlations. The Weight shift test showed a narrow range in scores (three to nine transfers) but the Motricity Index and Berg scales have a much larger range of scores (0-100 and 0-56 respectively). This difference in range of scores may account for the lack of correlation. Some subjects found it difficult to 'get the idea' of Weight Shift Test, especially if it was unfamiliar to them. The test was originally chosen as it was frequently used in clinical practice, but it may be that fashions in clinical practice have changed and it is less commonly or widely used. Given the relatively low validity of the Weight Shift test, there is an argument to remove it from the scale. It has been retained, however, as it has passed all the other stages of development (although at a lesser level that the other tests) and does show a relationship with disability (the Rivermead Mobility Index and Barthel Index) and handicap (the Extended Activities Of Daily Living Scale). However the clinical significance of changes in this performance test should be treated with more caution than the other performance tests.

The Nottingham Extended ADL scale only showed a significant correlation with the Weight Shift test. The poor correlation with the other stepping tests may be because of the small numbers involved, but could also be an indication that balance disability is not the main factor limiting motor handicap. Other factors such as lack of confidence or suitable transport may be as limiting, if not more so.

The Brunel Balance Assessment, the sub-scales and performance tests also showed good discriminant validity (except the Weight Shift test). The differences between the sub-scales were particularly marked with people in each sub-group showing very significant differences ($p<0.001$) in all outcome measures. The strength of the differences between the sub-groups (sitters, standers or steppers) suggests that this could be a way of stratifying or categorising patients for clinical or research purposes.

4.2.5.4 Conclusions

Concurrent, criterion-related and discriminant validity of the Brunel Balance Assessment, the sub-scales and the individual performance tests has been demonstrated. An exception is the Weight Shift test, which lacked discriminant validity. It has, however, passed all the other aspects of scale development (albeit at a lower level than the other performance tests) and so will be maintained as part of the scale, but the clinical significance of changes in the weight shift test should be treated with caution.

4.3 Conclusions to Chapter 4

The Brunel Balance Assessment has been developed, the psychometric properties have been successfully tested and the utility criteria have been met. Content and construct validity have been demonstrated in the development of the prototype. The resulting assessment forms a Guttman-type scale with the attendant advantages this offers. Although the scope of testing was quite limited, reliability has been shown for a wide range of severity and duration of hemiplegia when used by experienced physiotherapists and measurement error has been calculated for the scale, sub-scales and performance tests. Concurrent, criterion related and discriminant validity of the scale, sub-scales and performance tests have also been demonstrated by comparison with other tests of balance, motor impairment and disability and handicap. All indicate that the Brunel Balance Assessment is a valid measure of balance. Additional utility criteria have also been met: It is simple, quick and cheap to use, suitable for use in different settings, and a wide range of severity of hemiplegia, and it is clearly relevant to clinical practice.

In conclusion, the Brunel Balance Assessment has met all the criteria for use as an outcome measure to test balance in people with stroke. An indication of its suitability for clinical practice is that two of the participating clinical units are already using it as an outcome measure for their interventions.

Chapter 5

Characterising balance disability

5.1 Introduction

The previous study concerned the development of a new measure of balance disability that met the criteria for clinical utility. The original premise for this was that the rehabilitation of balance disability was a fundamental part of stroke physiotherapy, since balance is believed to be a basic requirement for other functions such as activities of daily living. Despite the logic of this belief, there has been little investigation of the relationship between balance ability and function, or how impairments contribute to balance disability. Previous literature has concentrated on identifying factors which, at admission, predict outcome in terms of activities of daily living, length of stay in hospital, or destination on discharge (see Kwakkel et al 1996 for a comprehensive systematic review). The aim of this study was to characterise balance disability using the Brunel Balance Assessment by investigating how impairments contribute to balance disability and investigating the relationship between balance disability and the activities of daily living. In doing this the following research questions were addressed:

- What profiles, in terms of impairments and demographics do people with different levels of balance ability display? What relationships exist between balance disability, impairments and demographics?
- Does balance make a significant contribution to ability in the activities of daily living?

5.2 Methods

Subjects

People with stroke were recruited from the participating hospitals using the recruitment methods and inclusion/exclusion criteria as described in Chapter 4, Section 4.2.1. The data used in this chapter were from the same subjects as were presented in the validity testing of chapter 4 (Section 4.2.5) and the data for this chapter and Section 4.2.5 were collected in one testing session. Each subject was visited once by a non-treating physiotherapist and testing took place in the physiotherapy treatment area or on the ward.

Factors associated with recovery from stroke were tested using the methods outlined in the 'Materials' section below. The procedure to carry out each of these tests is detailed in Appendix IV.

Materials

In order to address the research questions, valid ways of measuring the variables needed to be identified. The results of a previous comprehensive systematic review (Kwakkel et al 1996) identified a number of valid predictors of poor recovery following stroke and these were used as a basis for the choice of variables for this study. The predictors were old age, urinary incontinence, altered consciousness at onset, disorientation in time and place, severity of paralysis; sitting balance, low admission ADL and low levels of social support. Visuo-spatial defects and hemianopia were identified as negative predictors in some studies, but sex, side of stroke and ethnic origin were rejected as relevant predictors. Several other factors were considered important and clinically relevant but the lack of good reliable, valid, standardised outcome measures made detection of their influence on recovery less likely. These included sitting balance, spasticity, and proprioception. The type of stroke (based on clinical symptoms) has also recently been found to be a strong predictor of recovery in terms of ADL and walking ability (Sanchez-Blanco et al 1999). These factors were considered in the present study. Some were rejected for this study as they were controlled for by the admission criteria. All subjects were able to give informed consent, had a first-ever stoke and were previously functionally independent and mobile. This excluded people who were disorientated in time and place or unconscious, or who had previous history of stroke or limited ADL before admission. The subjects were also nearly all in-patients, which excluded social support as a factor. Urinary incontinence was rejected, as it is included in the Barthel Index. Similarly, sitting balance was included in the Brunel Balance Assessment. Spasticity was rejected for the lack of a suitable measurement tool (Pandyan et al 1999; Gregson et al 2000 & 1999; Pomeroy et al 2000), but proprioception was included using a recently published measurement tool (Winward et al 2000).

The list of factors used in the subsequent analysis and the measurement tools used to test them are shown in Table 5.1. Sex, side of stroke and time since stroke were also included for completeness.

Table 5.1. The basket of variables and the measurement tools used

Variable	Measurement tool	References
Balance ability	Brunel Balance Assessment	Chapter 4
Activities of daily living	Barthel Index (BI)	Collin et al 1988
Strength of the lower limb	Motricity Index (MI)	Demeurisse et al 1980
Visuo-spatial inattention	Star Cancellation Test	Wilson et al 1987 Halligan et al 1989
Hemianopia	Clinical confrontation test	Wade et al 1985
Sensation of the foot/ankle	Rivermead Assessment of Somatosensory Performance (RASP)	Winward et al 2000
Type of stroke	Stroke classification	Reding & Potes 1988
Age	Years	
Side of stroke	Left/ right	
Time since stroke	Weeks	
Sex	Male/ female	

Statistical Analysis

A variety of methods were used to address the research questions.

Defining and describing balance disability

The profiles (from the list of variables) of people with different levels of balance disability were examined. Balance disability was categorised into three levels of ability according to the score on the Brunel Balance Assessment; sitters (score 1-3); standers (score 4-6) and steppers (score 7-12). Differences between the groups were tested using the Kruskal-Wallis test. Relationships between the variables and balance ability were then explored using Pearson's correlations and the associations were examined further using multi-factorial linear regression. Each variable was regressed individually against balance disability using the 'enter' method. Then the factors that showed a significant association were entered together to examine the strength of association in combination.

The contribution of balance disability to every-day function

Variables previously known to be predictors of ADL ability were entered into a combined linear regression model. Balance disability was then added to the resulting model to see whether this affected the significant variables and made an improvement in the amount of variance explained. The variables suggested by Kwakkel et al (1996) as predictors of poor recovery of ADL: old age, severe paralysis, visuo-spatial defects and hemianopia, were included in this analysis. Type of stroke (Sanchez-Blanco 1999) and sensation were also included. Kwakkel et al (1996) had been unable to come to a conclusion about this variable for the lack of suitable outcome measures.

5.3 Results

55 people were recruited, and completed the testing. They were the same people as those included in the validity testing of Chapter 4 (section 4.2.5). There were 32 men (58%) and 23 women. 29 (53%) had a left hemiplegia, and 26 (47%) had a right hemiplegia. Their details are shown in Table 5.2 below.

Table 5.2 Details of the subjects included in Chapter 5

	Mean	Standard deviation	Range
Age (years)	68.3	12.8yrs	40-88
Time since stroke (weeks)	9 (median)	4,24 (IQR)	
Brunel	6.3	3.8	1-12
Berg	24	20-6	0-56
Rivermead	6.2	5	1-14
Motricity Index (leg)	51.8	30.2	0-92
Barthel Index	11.7	5.9	1-20
RASP	27.9	19.3	0-48
5m walk time (s)	10.0	7.1	4.4-22.5
Neglect	n=15 (27%)		
Hemianopia	n=16 (29%)		
Type of stroke	M-only=18 (32.7%) M-S =18 (32.7%) MSH=20 (36%)		

Brunel = Brunel Balance Assessment, Berg = Berg Balance Test, Rivermead = Rivermead Mobility Index, IQR = inter-quartile range, MI = Motricity Index, RASP = Rivermead Assessment of Sensori-motor Performance, M-only = motor-only type stroke, M-S =motor-sensory stroke, MSH = motor-sensory-hemianopic stroke.

5.3.2 Defining and describing balance disability

Profiles of people with different levels of balance disability

The profiles of the whole group and for people with different levels of balance ability (sitter, stander or stepper) are shown in Table 5.3. There were significant differences between groups in all variables except time since stroke, sex and side of stroke. The 'sitters' were most impaired and disabled, and the steppers least so.

Table 5.3 Profiles of people with different levels of balance disability.

	Total Group (n=55)	Sitters (n=16)	Standers (n=16)	Steppers (n=23)	p-value
Age (years)	68.3 (12.8)	74.8 (11)	67 (10.9)	64.9 (13.8)	0.027*
Time since stroke (weeks)	12.9 (13.8)	7.1 (5.6)	16.3 (16.9)	14.7 (14.8)	0.244
Sex (male)	32 (58%)	7 (44%)	10 (62.5)	14 (61%)	0.447
Side (left)	29 (53%)	9 (56%)	7 (44%)	14 (61%)	0.647
Type of stroke	M=18 MS=18 MSH=20	M=0 MS=6 MSH=10	M=4 MS=6 MSH=6	M=14 MS=5 MSH=4	0.001*
Barthel Index	11.7 (5.9)	5.1 (2.9)	9.9 (1.8)	17.4 (3.3)	0.001*
Motricity Index	51.8 (30.2)	23.6 (21.3)	41.3 (13.9)	77.6 (20.9)	0.001*
RASP	27.9 (19.3)	11.1 (15)	29.7 (17)	37.9 (16)	0.001*
Neglect (present)	15 (27%)	7 (44%)	6 (37.5%)	2 (8.7%)	0.004*
Hemianopia (present)	16 (29%)	10 (62.5%)	3 (8.8%)	3 (13%)	0.002*

M = motor-only type stroke, M-S =motor-sensory stroke, MSH = motor-sensory-hemianopic stroke, RASP = Rivermead Assessment of Sensori-motor Performance.

The trend between balance ability and the other variables was tested with Pearson's correlations. This showed significant relationships between balance disability (Brunel Balance Assessment score) and all the 'other variables' except time since stroke (Table 5.4). The negative value with age indicated that balance disability tended to get worse with increasing age. The nominal level variables (side of hemiplegia and sex) were not included in this calculation.

Table 5.4 Relationship between balance disability and other variables.

	Correlation coefficient	P value
Barthel Index	0.956	<0.0001
Weakness	0.865	<0.0001
Sensation	0.634	<0.0001
Neglect	0.509	<0.0001
Age	-0.312	<0.019
Time since stroke (weeks)	0.157	<0.272
Type of stroke	-0.613	<0.0001
Hemianopia	0.512	<0.0001

To test these associations further. each variable was individually entered into a linear regression model with balance disability as the dependent variable. Significant associations were found between balance disability and weakness, sensation, neglect, hemianopia, stroke type and age. Side of hemiplegia, time since stroke and sex were not associated with balance disability (Table 5.5).

Table 5.5. Multiple regression associations between balance disability and individual variables

	R-sq	β Coefficient	Significance	95% CI	
				Lower	Upper
Weakness	0.74	0.116	0.0001**	0.096	0.134
Sensation	0.4	0.132	0.0001**	0.086	0.174
Neglect	0.255	0.119	0.0001**	0.061	0.171
Hemianopia	0.257	0.507	0.0001**	2.354	6.508
Stroke type	0.387	-2.98	0.0001**	-4.008	-1.946
Age	0.101	-0.01	0.018*	-0.180	-0.018
Time	0.031	0.05	0.220	-.032	0.134
Sex	0.022	-1.33	0.285	-3.359	1.006
Side	0.002	0.187	0.730	-1.811	2.567

R-sq = the amount of variance in balance ability explained by the factor.
** = Significant at p <0.01 level. *= significant at p<0.05

The sum of the variances explained by the factors was more than 100% indicating some co-linearity between variables. To identify independent factors all the significant variables were entered into a combined linear regression model. The combined factors explained 83% of variance and only weakness, sensation and age remained as significant independent factors of which weakness was the most statistically significant (Table 5.6).

Table 5.6. Combined linear regression model of all the individually significant factors against balance disability.

	β Coefficient	Significance	95% Confidence Interval	
			Lower	Upper
Weakness *	0.660	0.0001	0.068	0.108
Sensation *	0.183	0.035	0.003	0.072
Age *	-0.161	0.013	-.089	-0.011
Stroke type	-0.141	0.193	-1.701	0.353
Hemianopia	0.051	0.648	-1.518	2.416
Neglect	-0.042	0.677	-.057	0.037

Does balance make a significant contribution to ADL ability?

Variables which had previously been found valid predictors of recovery of ADL were loaded into a combined linear regression model with ADL ability (Barthel Index) as the dependent variable. This explained 82.5% of variance and showed weakness and age as independent significant contributors to ADL ability (Table 5.7). Balance disability (Brunel Balance Assessment score) was then loaded into the model. The resulting model explained 93% of variance and balance disability and neglect were identified as significant independent variables (Table 5.8). This indicated that balance disability was a significant contributor to every-day function and its inclusion in the model improved the amount of variance explained.

Table 5.7 Linear regression model of ADL ability including all individually significant variables

	β Coefficient	Significance	95% Confidence Interval	
			Lower	Upper
Weakness *	0.672	0.000	0.103	0.161
Neglect	0.113	0.267	-0.031	0.108
Hemianopia	0.020	0.855	-2.646	3.179
Stroke type	-0.079	0.452	-2.077	0.939
Sensation	0.126	0.139	-0.013	0.090
Age *	-0.143	0.025	-0.125	-.009

* Significant at p<0.05

Table 5.8. Linear regression model of ADL ability with balance disability included.

	β Coefficient	Significance	95% Confidence Interval	
			Lower	Upper
Balance *	0.786	0.000	0.883	1.431
Weakness	0.146	0.067	-0.002	0.059
Sensation	-0.018	0.746	-0.040	0.029
Neglect *	0.146	0.028	0.006	0.095
Hemianopia	-0.026	0.719	-2.202	1.530
Stroke Type	0.019	0.781	-0.842	1.114
Age	-0.021	0.624	-0.049	0.030

* Significant at p<0.05

5.4 Discussion

In this chapter balance disability has been described and profiles of people with different levels of disability studied. The importance of the various factors to balance disability has been investigated and the importance of balance disability to everyday function tested. The results demonstrated that people with different levels of balance are heterogeneous and show very different profiles in the type and extent of impairments and disabilities. Not surprisingly poor balance ability was associated with severity of weakness and sensory loss, presence of hemianopia and neglect, extent of the lesion and increasing old age, and with greater disability.

The strength of the difference in profiles between people with different levels of ability (sitters, standers and steppers) indicates that people with different balance abilities should be treated as different groups - intervention and recovery studies of motor abilities for example. The results suggest that the Brunel Balance Assessment can be used to stratify balance disability. Many previous studies have identified different recovery patterns post-stroke (e.g. Skilbeck et al 1983; Wade & Langton-Hewer 1987; Horgan & Finn 1997; Sanchez-Blanco et al 1999; Smith & Baer 1999 for example) but no way of defining groups with different recovery profiles has been commonly accepted. Balance disability has been found to make a strong contribution to ADL ability over and above previously known contributors. The inclusion of Brunel Balance Assessment, possibly in a basket of variables such as type of stroke and time since stroke may enable such groups to be defined. Further study would be needed to test this idea.

Which factors contribute to balance disability?

The results of this study found that weakness (of the leg), sensation and age were significant independent contributors to balance disability. Three other studies considering the relationship between balance and other variables were found.

The most closely aligned study was by Keenan et al (1984) who investigated factors affecting balance and walking post stroke in 90 stroke patients using the Rancho Los Amigos Assessment. Significant correlations were found between balance and motor control, proprioception of the leg, muscle tone, body awareness and sensory integration and but no relationship was found between balance and age, time since stroke, visual field deficits, tactile sensation, and ability to follow commands. Comparisons between the two studies are hampered by the lack of definition of several of the factors used by Keenan et al (1984). In particular body awareness, visual field deficits, ability to follow commands and sensory integration are undefined. But the finding that motor control and some aspects of sensation are correlated to balance ability supports the results reported here. The lack of relationship with age and balance contrasts with the results of this study but may be explained by the different inclusion criteria. Keenan et al included people of any age and the subject group ranged from 22 to 85 years with a mean of 59 years, in contrast this study excluded people under forty years of age and had a mean age of 68 years.

Bohannon (1989) investigated the relationships between static standing balance and strength of the leg, age, sex, side of stroke and time since stroke in a group of 33 acute patients. Significant correlations between standing balance and lower limb strength were found but not with balance and sex, time since stroke or age. The age of the subjects are similar in both studies and the correlation coefficients (both used Spearman's rho) are similar (0.367 for this study and 0.409 on admission for Bohannon).

Further support for a relationship between sensation and balance was found by Niam et al (1999). They studied the relationships between postural sway and the Berg Balance test in a convenience sample of 30 ambulant chronic hemiplegics. The factors relevant to this study were proprioception of the ankle, age, side of hemiplegia and time since stroke. Age and time since stroke was not related to postural sway or balance score, people with impaired ankle proprioception had poorer postural sway and balance. Interestingly, people with a left hemiplegia tended to show greater postural sway than people with a right hemiplegia, although this difference did not extend to a difference in the Berg Balance test. Niam et al's subject group were younger (mean age 58 years) and more able (they were all independently mobile) than in the present study which may explain the difference in relationship with age.

These other studies support the findings of the present study in that weakness and sensation are significant contributors to balance disability. They also support the lack of significance of sex, side of stroke and time since stroke. The significance of age as a factor in balance ability post-stroke is more controversial, although difference in inclusion criteria may explain the contradictory findings. The importance of age in balance disability post-stroke is complex. Balance and postural control tend to decline with increasing years (Lord & Hall 1994), older people tend to have more severe strokes (Carlo et al 1999; Arboix et al 2000) and increasing age is a negative predictor of recovery from stroke (Kwakkel et al 1996), so it is not surprising that balance disability tends to be greater in older people post-stroke. However whether these negative findings are due to the effects of aging per se, or are due to the complicating co-pathologies and pre-existing disabilities, which are more frequent with increasing years, is unclear. This study attempted to exclude people with co-morbidities but it may be that subjects had sub-acute or mild pathologies which were not excluded in the selection criteria.

The effects of visual field defects and visuo-spatial neglect did not emerge as a significant factor in balance disability which was surprising given the negative impact that the presence of neglect has on level of disability and rehabilitation outcome (Paolucci et al 2001).

Time since stroke also failed to emerge as a factor contributing to balance disability. This was a surprise as balance and mobility generally improve over time (Partridge et al 1987; Smith & Baer 1999), so acute subjects may have been expected to have poorer balance. However, the subject group represented a convenience sample of people receiving physiotherapy rather than a cross-section of all people with stroke. A selection bias existed, as people who recovered well would be discharged from physiotherapy and therefore not be recruited, with the result that the chronic subjects that were recruited were people with more severe problems who continued to receive physiotherapy. Similarly, recruitment of the most severely impaired acute subjects was limited, if they were unable to give consent or were too unwell to participate. Consequently, recruitment of the most and least able subjects at any time since stroke was limited. As the aim of this study was to use the Brunel Balance Assessment with people receiving physiotherapy, this selection bias was inevitable. Unfortunately no information was gathered about people who were not recruited to the study so further analysis is not possible.

5.4.1 The importance of balance for every-day activities

The finding that balance was the strongest contributor to ADL ability adds support to the clinical belief that balance is a fundamental prerequisite for other functions (Carr & Shepherd 1987, Lennon & Ashburn 2000) and supports the clinical usefulness of a measurement tool that specifically assesses balance ability. It is also concordant with previous authors who found that balance is a predictor of recovery. The absence of static sitting balance on admission is a strong negative predictor of recovery in ADL and walking (Keenan et al 1984; Wade & Hewer 1987; Loewen & Anderson 1990; Morgan 1994; Feigin et al 1996; Franchignoni et al 1997). Conversely 'high postural stability' at admission to rehabilitation is a strong predictor of improvement in ADL during rehabilitation (Lofgren et al 1998).

It is interesting that balance was a stronger contributor than other variables such as weakness, sensory loss or neglect which were considered important in recovery of ADL (Kwakkel et al 1996). Lofgren et al (1998) also found co-linearity between postural stability (as measured by the balance section of the Fugl-Meyer scale) and measures of motor function such as degree of active movement, range of movement, sensation and mobility and found postural stability to be the strongest factor. It may be that balance ability represents a 'final common movement pathway' in the same way that urinary continence has been described as a final common ADL pathway (Tallis 2001, personal communication).

Limitations of the study

Several limitations need to be born in mind when considering these results. Firstly, the selection bias. The aim of the study was to recruit people with balance disabilities due to a stroke who represented a cross-section of people receiving physiotherapy. This was achieved but at a cost of the generalisability of the results. Many people with stroke who may have influenced the results presented here were excluded. Recruitment of the most severely affected stroke patients was limited by excluding people who were not able to give consent or medically unstable. Excluding people with pre-existing ADL or mobility problems also meant many people receiving physiotherapy were not recruited.

The choice of independent variables would also affect the results. In particular muscle tone and spasticity were not tested, because of the lack of suitable measurement tools, but physiotherapists attach great importance to these factors (Lennon & Ashburn 2000). If these factors had been included the results may have been different. Other variables which may have influenced the results but not include were measures of cognitive function, visual acuity and perceptual skills beyond the most basic test of visuo-spatial neglect. Lack of suitable outcome measures precluded their use but further research would need to consider them. Nevertheless the high degree of variance explained in the regression models suggests that most important variables were included.

A final factor was the small number of subjects. The numbers recruited were based on a recommendation by Kwakkel et al (1996) that ten subjects should be recruited for each variable. However, after data collection had been completed another text recommended that at least 50 subjects plus eight for each variable should be recruited (Tabachnick & Fidell 1996). If this formula had been used more than twice as many subjects should have been recruited.

Further research is needed with larger cohorts, wider selection criteria to including as many people as possible over a defined time-period, and measuring other variables such as muscle tone and attention so that the results can be more widely generalized and applied to the clinical setting.

Finally, this study is unusual in that it has predicted disability based on the number and severity of impairments and other demographic details. It has not predicted recovery or outcome which are the more usual dependent variables for prediction studies (see Kwakkel et al 1996). This was because the aim of the study was to characterise balance disability, rather than explore the recovery of balance. Recovery patterns of balance for people with different levels initial loss and how this could predict eventual outcome remains to be done.

5.5 Conclusion

Balance disability in people with stroke was found to be heterogeneous. People could be divided into different groups according to their performance in the Brunel Balance Assessment. Consequently the Brunel Balance Assessment could be used as a stratification tool for research studies considering balance disability. Balance was found to be a strong contributor to independence in every-day activities, confirming the clinical relevance of its measurement. Many factors may impact on balance ability but in this sample weakness, sensation and age were the most significant contributors.

Chapter 6

Conclusions

The aim of the present study was to address some of the issues limiting research into stroke physiotherapy and the development of clinical practice. In Chapter One the main problems were identified as a lack of models to describe and explain clinical practice, lack of suitable outcome measures and lack of methods to stratify people with stroke to limit the inherent variability found in such a heterogeneous group. These problems have been addressed with respect to balance in people with stroke. This aspect of the physical consequences of stroke was chosen, because although its rehabilitation is considered a cornerstone of physiotherapy (Ballinger et al 1999; Lennon & Ashburn 2000), there has been relatively little research in comparison to other activities such as walking or upper limb function.

In a comprehensive review, the positive and negative aspects of current measurement tools were identified. None met all the utility criteria but the 'best of the bunch ' were identified. The main outcome of this series of studies was to develop an outcome measure that did meet all of the utility criteria.

6.1 The model of clinical practice

A model of the clinical process of assessing posture and balance was developed from reviews of experiential literature and consultations with neurological physiotherapists. Although not a traditional focus group method, the methodology used to consult physiotherapists about their practice drew on this technique and addressed the aim of the research and a model of clinical practice was developed. The key reasoning processes, subjective measurements and decisions physiotherapists used while assessing posture and balance in people with stroke were identified and this informed the content of the new outcome measure. The method and basic model developed here could be used to develop models for assessment of other aspects of physiotherapy, such as mobility and upper limb function. Although the 'important physical features' would differ with other aspects of physiotherapy, the clinical reasoning process may be essentially the same. The methodology could also be used to develop models of treatment and may also be generalisable to other neurological conditions.

Any model of practice needs to be reflective of current practice and this is known to change and develop with time (Lennon & Ashburn 2000), so physiotherapists' practice would need to monitored periodically to ensure that the model still reflected current practice and adapted if necessary. Some variations in practice have been identified based on the physiotherapists' geographical location and the clinical area in which they worked (Davidson & Waters 2000; Lennon & Ashburn 2000). This model has been presented to physiotherapists working throughout the UK and was felt to reflect their practice, so it would appear to have overcome the possible geographic variation, although further study with more rigorous and formal assessment methods could confirm this. The physiotherapists in this study worked in stroke, neurological elderly care and community services, so the model appears to overcome differences between different types of service. Again further study with more rigorous assessment methods is needed to confirm this.

Despite these limitations, the model is one of few attempts to define and describe the complex clinical reasoning processes that take place during stroke physiotherapy and is the first to deal specifically with the assessment process. The lack of explicit models and protocols to define, describe and explain physiotherapy practice has inhibited research and the development of clinical practice for years. This model fills part of that gap with respect to the assessment of balance and posture and offers a method for future research to address other areas of neurological physiotherapy.

6.2 The Brunel Balance Assessment

In Chapter 4, the development of a new outcome measure for balance drawing on the results of Chapters 2 and 3 was described. This combined ordinal scales with functional performance tests of balance. Unfortunately, no equipment that could assess posture in people with stroke (particularly those with a severe hemiplegia) and fulfil the utility criteria was found. For this, new equipment would need to be developed. This was beyond the scope of this project and has not been taken any further, although it is a matter for future research.

The prototype Brunel Balance Assessment was developed and the psychometric properties tested. It was found to form a Guttman-type hierarchical scale with strong internal consistency as a test of balance. It was reliable for use with people with a wide range of severity and duration of hemiplegia, and a wide range of ages (40+ years). All the testers were experienced neurological physiotherapists. The Brunel Balance Assessment cannot be assumed to be as reliable with less experienced physiotherapists or other health care professions. However, the judgements required do not require any professional judgement. They are essentially counting, measuring or timing so it is unlikely that the reliability would be significantly different with less experienced physiotherapists or other health care professions- further testing should establish this one way or the other.

Responsiveness to change was established by assessing the measurement error for the ordinal scale and each performance test. This was simple to calculate and since the data is in the same units as the measurements, it is clinically meaningful. The wider use of this method is recommended. The calculations were performed on data from a wide range of stroke patients, so the use of the measurement error to monitor clinical change is not restricted to any particular groups of patient. Inevitably, the numbers of subjects at each end of the scale of abilities were relatively small, and further research could examine measurement error of the sitting and stepping tests with larger numbers of subjects.

Finally, the validity as a measure of balance disability was assessed. Validity can never be said to be absolutely proven, but as far as is possible the Brunel Assessment has demonstrated validity as a measure of balance. Content and construct validity were developed from Chapter 3, while concurrent, criterion-related and discriminate validity and internal consistency was demonstrated in Chapter 4. Predictive validity has not been explicitly tested but some predictive value is drawn from the hierarchy and Guttman scaling. Further research could establish whether scores of the Brunel Assessment could predict abilities on discharge from rehabilitation or after a set time period. Another interesting issue for further research is to use the Brunel Balance Assessment to study patterns of recovery of balance ability in different sub-groups of patients.

The Brunel Balance Assessment has been developed to fulfil all the important utility criteria. It fills a gap in the range of available outcome measures and can be used for research or clinical purposes. It could also be used to assess the effects of physiotherapy interventions on balance in people with stroke.

6.3 Is balance important?

The original premise of this thesis was that balance is an essential aspect of physiotherapy for people with stroke but that research had been hampered by the lack of suitable outcome measures. In Chapter 5, the Brunel Balance Assessment was used to test this clinical belief. The evidence supported this belief. Balance was found to be the most statistically significant contributor to ADL ability, while lower limb strength, sensation and age were the most significant contributors to balance disability. The findings add to the body of knowledge about the balance disability and its inter-relationships with general function and impairments in people with stroke.

People with different levels of balance disability were noted to have markedly differing profiles in the number and severity of abilities and impairments. Levels of balance can easily be stratified into sitting, standing or stepping balance using the Brunel Balance Assessment. This is recommended as a simple, valid and reliable way to stratify and characterise severity of the physical consequences of stroke.

6.4 Limitations of the studies

A limiting factor throughout this series of studies has been the resources available to the author. This has influenced the methodology adopted in many ways. The main limitation is the small numbers recruited to all the studies. There are few guidelines or 'rules of thumb' available to calculate population sizes for psychometric testing studies and so sample sizes were decided pragmatically based on the number of subjects that could be found in a limited time span with one person collecting the data. The financial and time resources influenced those decisions. The results of these studies need to be replicated in larger studies to improve generalisation and robustness.

There are two areas where the small number had a particular impact. Firstly the testing of reliability and responsiveness to change was limited in scope. A stronger indication of the inter-tester reliability and error would have been gained if more physiotherapists had been included. This information is needed before the

Brunel Balance Assessment can be used in the clinical setting with confidence. The involvement of more subjects at the extreme ends of ability would have also made the results more robust. Secondly, the small numbers affected the regression analysis in Chapter 5. With a larger sample, more significant relationships may be found. However, the fact that significant relationships were found even with a small sample size gives an indication of the strength of the relationships.

A second limitation was the effect of the selection criteria. The overall aim was to recruit subjects to represent people with stroke who received physiotherapy, so that the Brunel Balance Assessment would be generalisable. However, methodological and ethical considerations limited this the recruitment. People with balance disabilities from other pathologies or previous strokes and people who could not give informed consent were excluded. This means that the Brunel Balance Assessment has not been tested on a group stroke patients with the most severe disabilities. Some people with severe disabilities from their stroke and who could give consent were included, but the robustness of generalisation to people with severe disabilities is limited.

Finally, the Brunel Balance Assessment has not been used to measure recovery or changes over time. It could be argued that this is not necessary in the development of a new measurement tool and the study to characterise balance disability. It was also influenced by pragmatic resource-related difficulties in undertaking a longitudinal study. However whether the Brunel Balance Assessment can be used to predict recovery and outcome remains as a question for further study.

Final Conclusion

This research has fulfilled its aim to address the main factors limiting the development of the evidence-base for stroke physiotherapy with reference to balance. In doing so, it has offered a new and unique contribution to the body of knowledge about balance disability following stroke by:

- developing a clinical model to define and describe an aspect current practice namely assessment of posture and balance

- developing a robust new outcome measure which addresses all the criteria for use in the research or clinical setting
- suggesting a simple, reliable, valid method of stratifying people with stroke, and characterising the physical consequences of stroke
- exploring the role and importance of balance within stroke disability using the Brunel Balance Assessment.

References

Adams SA, Ashburn A, Pickering R, & Taylor D (1997a). The scalability of the Rivermead motor assessment in acute stroke patients. Clinical Rehabilitation 11;42-51

Adams SA, Pickering R, Ashburn A & Lincoln N (1997b). The scalability of the Rivermead motor assessment in non-acute stroke patients. Clinical Rehabilitation 11;52-59

Altman D & Bland M (1983). Measurement in medicine; the analysis of method comparison studies. The Statistician 32;307-317.

Andrews K, Brocklehurst J, Richards B and Laycock P (1982). The recovery of the severely disabled stroke patient. Rheumatology and Rehabilitation 21;225-230

Arboix A, Garcia-Eroles L, Massons J, Oliveres M, Targa C (2000). Acute stroke in very old people; Clinical features and predictors of in-hospital mortality. Journal of American Geriatric Society 48;1;36-41.

Ashburn A. A physical assessment for stroke patients (1982). Physiotherapy 68;4;109-113.

Ashburn A, Partridge C, De Souza L (1993). Physiotherapy in the rehabilitation of stroke: A review. Clinical Rehabilitation 7;337-45.

Ashburn A (1995). A review of current physiotherapy in the management of stroke. In Physiotherapy in stroke management (ed Harrison MA), Churchill Livingstone, Edinburgh.

Ashburn A (1996a). Sitting balance responses; Comparison of subjects with hemiplegia and normal controls. Proceedings of Physiotherapy Research Society Physiotherapy 82;11;636.

Ashburn A (1996b). A study of normal sitting balance responses. Proceedings of Physiotherapy Research Society Physiotherapy 82;11;636.

Atkinson G & Neville A (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Medicine Oct;26;4 ;217-238

Badke M & Duncan P (1983). Patterns of rapid motor responses during postural adjustments when standing in healthy subjects and hemiplegic patients. Physical Therapy 63;13-20.

Ballinger C, Ashburn A, Low J & Roderick P (1999). Unpacking the black box of therapy - a pilot study to describe occupational therapy and physiotherapy interventions for people with.stroke. Clinical Rehabilitation 13;301-309.

Bamford J, Sandercock P, Dennis D, Burn J & Warlow C (1988). A prospective study of acute cerebrovascular disease in the community; The Oxfordshire Community Stroke project. 1: Methodology, demography, and incident cases of first-ever stroke. Journal of Neurology, Neurosurgery and Psychiatry 51;1373-80.

Bamford J, Sandercock P, Dennis D, Burn J & Warlow C (1990). A prospective study of acute cerebrovascular disease in the community; The Oxfordshire Community Stroke project 1981-1986. 2: Incidence, case fatality rates and overall outcome at one year. Journal of Neurology, Neurosurgery and Psychiatry 53;16-22

Bamford J, Sandercock P, Dennis M, Burn J, Wardlow C (1991). Classification and natural history of clinically identifiable subtypes of cerebral infarction. Lancet 337;1521-26.

Barer D. & Nouri F. Measurement of activities of daily living (1989). Clinical Rehabilitation 3;179-87.

Basmajian JV, Gowland CA, Finlayson M (1987). Stroke treatment; comparison of integrated behavioural physical therapy versus traditional physical therapy programs. Archives of Physical Medicine & Rehabilitation 68;267-72.

Beckerman H, Vogelaar T, Lankhorst G, Verbeek A (1996). A criterion for stability of motor function of the lower extremity in stroke patients using the Fugel-Meyer Assessment. Scandinavian Journal of Rehabilitation Medicine 28;3-7

Benaim C, Perennou D, Villy J, Rousseaux M, Pelissier J (1999). Validation of a standardised assessment of postural control in stroke patients: The Postural Assessment Scale for Stroke Patients (PASS). Stroke 30;1862-1868.

Berg K, Wood-Dauphinee S, Williams J & Gayton D (1989). Measuring balance in the elderly; preliminary development of an instrument. Physiotherapy Canada 41;6;304-3411

Berg K, Wood-Dauphinee S, Williams J & Maki B (1992a). Measuring balance in the elderly: Validation of an instrument. Canadian Journal of Public Health 83;supp 2;S7-11

Berg K, Maki B, Williams J, Holliday P & Wood-Dauphinee S (1992b). Clinical and Laboratory measures of postural balance in an elderly population. Archives of Physical Medicine and Rehabilitation 73;1073-1081

Berg K, Wood-Dauphinee S, Williams J (1995).The Balance Scale; Reliability assessment with elderly residents and patients with acute stroke. Scandinavian Journal of Rehabilitation Medicine 27;27-36

Bernhardt J, Ellis P, Denisenko S & Hill K (1998). Changes in balance and locomotion measures during rehabilitation following stroke. Physiotherapy Research International 3;2;109-122.

Bland M (1995). An introduction into medical statistics (2nd ed). Oxford University Press

Bobath B. (1990) Adult hemiplegia; Evaluation and treatment, 3rd ed. Butterworth-Heinnemann, Oxford

Bohannnon R, Larkin P, Cook A, Gear J, Singer J (1984). Decrease in timed balance scores. Physical Therapy 64;1067-1070.

Bohannon R, Smith M & Larkin P (1986). Relationship between independent sitting balance and side of hemiplegia. Physical Therapy 66;6;944-5

Bohannon R (1987). Gait performance of hemiparetic stroke patients; Selected variables. Archives of Physical Medicine and Rehabilitation 68;777-781.

Bohannon R & Smith MB (1987). Inter-rater reliability of the modified Ashworth scale of muscle spasticity. Physical Therapy 67;206-7.

Bohannon R (1989). Correlation of lower limb strengths and other variables with standing performance in stroke patients Physiotherapy Canada 41;4;198-202.

Bohannon R & Waldron R (1991). Weight-bearing during comfortable stance in patients with stroke: Accuracy and reliability measurements. Australian Physiotherapy 37;1;19-22. .

Bohannon R (1992). Lateral trunk flexion strength; impairment, measurement reliability and implications following unilateral brain lesions. International Journal of Rehabilitation Research 15;249-251.

Bohannon R, Walsh S & Joseph M (1993). Ordinal and timed balance measurements; reliability and validity in stroke patients. Clinical Rehabilitation 7;9-13.

Bohannon R (1995). Standing balance, lower extremity muscle strength and walking performance of patients referred for physical therapy. Perceptual and Motor Skills 80;379-385.

Bohannon R. & Leary K (1995). Standing balance and function over the course of acute rehabilitation. Archives of Physical Medicine & Rehabilitation 76;944-6.

Bowling A. (1997). Research Methods in Health: Investigating Health and Health Services. Open University Press, UK.

Brauer S, Burns Y & Galley P (1999). Lateral Reach: A clinical measure of medio-lateral postural stability. Physiotherapy Research International 4;2;81-88.

Briggs R, Gossman M, Drews J Shaddeau S (1989). Balance performance among non-institutionalised elderly women. Physical Therapy 69;748-756.

Brocklehurst J, Robertson D, James-Groom P (1982). Clinical correlates of sway in old age - sensory modalities. Age & Ageing 11;1-10.

Bronstein A, Brandt T & Woolacott M (1996). Clinical disorders of balance, posture and gait. Arnold Pub. London

Brunnstrom S (1966). Motor testing procedure in hemiplegia based on sequential recovery stages. Physical Therapy 46;357-75.

Burdett R, Brown K, & Fall M (1986). Reliability and validity of four instruments for measuring lumbar spine and pelvic positions. Physical Therapy 66;5;677-684.

Burton A (1986). Regional lumbar sagittal mobility; measurement by flexi-curves. Clinical Biomechanics 1;20-26.

Caldwell C, MacDonald D, MacNeil K, McFArland K, Turnbull G & Wall J (1986). Symmetry of weight distribution in normals and stroke patients using digital scales. Physiotherapy Practice 2;109-116.

Carr E, Keeney F, Wilson-Barret J, & Newham D (1999). Inter-rater reliability of postural observation after stroke. Clinical Rehabilitation 13;229-242.

Carr J & Shepherd R (1980). Physiotherapy for Disorders of the Brain. Heinneman Medical, London

Carr J, Shepherd R, Nordholm R & Lynne D (1985). Investigation of a new motor assessment scale for stroke patients. Physical Therapy 65;2;175-179

Carr J, Mungovan S, Shepherd R, Dean C, & Nordholm L (1994). Physiotherapy in stroke rehabilitation: Bases for Australian physiotherapists 'choice of treatment. Physiotherapy Theory and Practice 10;201-209

Carr J & Shepherd R (1998). Neurological Rehabilitation: Optimising Human Performance. Butterworth Heinemann, London

Chandler J, Duncan P, Studenski S (1990). Balance performance on the Postural Stress Test; Comparison of young adults, healthy elders and fallers. Physical Therapy 70;410-415.

Chiarello C & Savidge R (1993). Inter-rater reliability of the Cybex EDI-320 and fluid goniometer in normals and patients with low back pain. Archives of Physical Medicine and Rehabilitation 74;32-7.

Cifu D & Stewart D (1999). Factors effecting functional outcome after stroke; A critical review of rehabilitation interventions. Archives of Physical Medicine and Rehabilitation 80;535-39.

Ciprany-Dacko L, Innerst D, Johannsen J & Rude V (1997). Inter-rater reliability of the Tinetti balance score in novice and experienced physical therapy clinicians. Archives of Physical Medicine and Rehabilitation 78;1160-1164.

Cohen J (1960). A co-efficient of agreement for nominal scales. Educational and Psychological Measurement 20;37-46.

Cohen J (1968). Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. Psychological Bulletin 70;213-20.

Cohen H, Blatchlly C & Gombash L (1993). A study of the clinical test of sensory integration and balance. Physical Therapy 73;6;346-351.

Collen F, Wade DT, Bradshaw C (1990). Mobility after stroke; reliability of measures of impairment and disability. International Disability Studies 12;6-9.

Collen F, Wade D, Robb G & Bradshaw C (1991). The Rivermead mobility index; a further development of the Rivermead motor assessment. International Disability Studies 13;50-54.

Collin C & Wade D. Assessing motor impairment after stroke: A pilot reliability study. Journal of Neurology, Neurosurgery & Psychiatry (1990);53;576-579

Collin C, Wade DT, Davis S, & Horne V (1988). The Barthel ADL Index: A reliability study. International Disability Studies 10;61-3.

Corriveau H, Hebert R, Prince F, Raiche M (2000). Intra-session reliability of the Centre of Pressure minus Centre of Mass variable of postural control in the healthy elderly. Archives of Physical Medicine and Rehabilitation 81;45-8.

Cotton E & Kinsman R (1987) Conductive Education for Adult Hemiplegia. Churchill Livingstone, London

Daley K, Mayo N, Danys I, Cabot R, Wood-Dauphinee S (1997). The Stroke Rehabilitation Assessment of Movement (STREAM); refining and validating the content. Physiotherapy Canada 49;4;4269-78

Daley K, Mayo N, Wood-Dauphinee S (1999). Reliability of scores on the Stroke Rehabiltiation Assessment of Movement (STREAM) measure. Physical Therapy 79;1;8-23.

Davidson I & Waters K (2000). Physiotherapists working with stroke patients: A national survey. Physiotherapy 86;2;69-80.

Davies P (1985). Steps to Follow. Springer-Verlag London

Davies P (1990). Right in the middle. Springer-Verlag London

Davies P (1994). Starting Again. Springer-Verlag London

Dean C & Mackey F (1992). Motor assessment scale scores as a measurement of rehabilitation outcome following stroke. Australian Journal of Physiotherapy 38;1;31-35.

Dean C & Shepherd R (1997). Task related training improves performance of seated reaching tasks after stroke; A randomised controlled trial. Stroke 28;722-28.

Demeurrisse G, Demol O & Robaye E (1980). Motor evaluation in vascular hemiplegia. European Neurology 19;382-9.

Department of Health (2000) The New NHS Plan HMSO, London

Depoy E & Gitlin L (1994). An introduction to research: Multiple strategies for health and human services. Mosby & Co. London

Dettmann M, Linder M, Sepic S (1987). Relationships between walking performance, postural stability and functional assessments of the hemiplegic patient. American Journal of Physical Medicine 66;2;77-90.

Di Carlo A, Lamassa M, Pracucci G, Basile A, Trefoloni G, Vanni P, Wolfe C, Tilling K, Ebrahim S, Inzitari D, for the European BIOMED study of stroke care group (1999). Stroke in the very old: Clinical presentation and determinants of 3-month functional outcome; A European perspective. Stroke 30;2313-2319.

Dickstein R & Dvir Z (1993). Quantitative evaluation of stance balance performance in the clinic using a novel measuring device. Physiotherapy Canada 45;2;102-108.

Dickstein R, Dvir Z, Jehousa E, Rois M, Pillar T (1994). Automatic and voluntary lateral weight shifts in rehabilitation of hemiplegic patients. Clinical Rehabilitation 8;91-99.

Dickstein D, Abulaffio N, Gelernter I & Pillar T (1996). An ultrasonic-operated kinematic measurement system for assessment of stance balance in the clinic. Clinical Biomechanics 11;3;173-175.

Dickstein R, Hocherman S, Pillar T & Shaham R (1996). Stroke rehabilitation; three exercise therapy approaches. Physical Therapy 66;1233-8.

DiFabio R & Badke M (1990). Relationship of sensory organisation to balance function in patients with hemiplegia. Physical Therapy 70;542-548.

Di Fabio R & Badke M (1990). Extraneous movement associated with hemiplegic postural sway during dynamic goal-directed weight redistribution. Archives of Physical Medicine and Rehabilitation 71;365-371.

Dodds T, Martin DP, Stolov W, Deyo RA (1993). Validation of the Functional Independent Measure and its performance amongst rehabilitation in-patients. Archives of Physical Medicine & Rehabilitation 74;531-6.

Donahue M, Riddle D, & Sullivan M (1996). Inter-tester reliability of a modified version of McKensie's lateral shift assessments obtained on patients with low back pain. Physical Therapy 76;7;706-716.

Dornan J, Fernie G, Holliday P (1978). Visual input: Its importance in the control of postural sway. Archives of Physical Medicine and Rehabilitation 59;586-591

Duncan P, Propst M & Nelson S (1983). Reliability of the Fugel-Meyer Assessment of Sensori-motor recovery following cerebro-vascular accident. Physical Therapy 63;10;1606-1611

Duncan P, Studenski S, Chandler J, Bloomfeld R & LaPointe L (1990). Electromyographic analysis of postural adjustments in two methods of balance testing. Physical Therapy 70;88-96.

Duncan PW, Weiner DK, Chandler J et al (1990). Functional Reach; A new clinical measure of balance. Journal of Gerontology 45;M192-M197.

Duncan P, Studenski S, Chandler J & Prescott B (1992). Functional reach: Predictive validity in a sample of elderly veterans. Journal of Gerontology 47;M93-8.

Duncan P, Goldstein L, Matchar D, Divine G Feussner J (1992). Measurement of motor recovery after stroke; Outcome assessment and sample size requirements. Stroke 23;1084-1089.

Duncan P. Richards L, Wallace D (1998). A randomised controlled pilot trial of home-based exercise program for individuals with mild or moderate stroke. Stroke 29;2055-60.

Dyer C (1995). Beginning research in psychology; A practical guide to research methods. Blackwell Publishers

Eakin P (1989). Assessment of activities of daily living; a critical review. British Journal of Occupational Therapy 52;1;11-15.

Early Supported Discharge Trialists (1999). Services for reducing duration of hospital care for acute stroke patients (Cochrane -Review). In the Cochrane Library, Issue 3, 1999, Oxford.

Edwards S. (1996) Neurological Physiotherapy: A Problem-Solving Approach. Churchill Livingstone, London

Effective Health Care Bulletin No. 2. (1992) Stroke Rehabilitation. Leeds University Printing Services

Elliot C & Murray A (1998). Repeatability of body sway measurements: day-to-day variation measured by sway magnetometry. Physiological Measurement 19;159-164.

Ernst E (1990). A review of stroke rehabilitation and physiotherapy. Stroke 21;1081-1085.

Evans M, Goldie P, Hill K (1997). Systematic and random error in repeated measurements of temporal and distance parameters of gait after stroke. Archives of physical Medicine and Rehabilitation 78;725-9.

Feigin L, Sharon B, Czaczkes B, Rosin A (1996). Sitting equilibrium two weeks after a stroke can predict walking ability at six months. Gerontology 42;348-353.

Feigenson J, McDowell F, Meese P, Greenberg S (1977). Factors influencing outcome and length of stay in a stroke rehabilitation unit. Stroke 8;6;651-656.

Fishman MN, Colby LA, Sachs LA, Nichols DS (1997). Comparison of upper extremity balance tests and force platform testing in persons with hemiparesis. Physical Therapy 77;10;1052-1062.

Fitzgerald JE, Murray A, Elliot C & Birchall J (1993). Comparison of balance assessment by sway magnetometry and force platforms. Archives of Otolaryngol Head and Neck Surgery 119;41-6.

Fitzgerald G, Wynveen K, Rheault W & Rothschild B (1983). Objective assessment with establishment of normal values for lumbar spine range of motion. Physical Therapy 63;11;1776-1781.

Forster A, Young J & Langhorne P (1996). Systematic review of day hospital care for elderly people; The Day Hospital Group. British Medical Journal 318;837-41.

Franchignoni F, Tesio L, Ricupero C & Martino M (1997). Trunk Control Test as an early predictor of stroke rehabilitation outcome. Stroke 28;1382-1385.

Frost M, Stuckley D, Smalley A, Dorman G (1962). Reliability of measuring trunk motions in centimetres. Physical Therapy 10;1431-1437.

Fugel-Meyer A, Jaakso L, Leyman I (1975). The post-stroke hemiplegic patient; A method for evaluation for physical performance. Scandinavian Journal of Rehabilitation Medicine 7;13-31

Gabell A & Simons M (1982). Balance coding. Physiotherapy 68;9;286-289

Gajdosik R, Simpson R, Smith R, Dontigny R (1985). Pelvic tilt; intra-tester reliability of measuring the standing position and range of motion. Physical Therapy 65;2;169-174.

Geurts A, Nienhuis B, Mulder T (1993). Intra-subject variability of selected force platform parameters in the quantification of postural control. Archives of Physical Medicine and Rehabilitation 74;1144-1150.

Geurts A, Ribbers G, Knoop J, & Limbeck J (1993). Identification of static and dynamic postural instability following traumatic brain injury. Archives of Physical Medicine and Rehabilitation 77;639-44.

Gill K, Krag M, Johnson G, Haugh L, Pope M (1988). Repeatability of four clinical methods for assessment of lumbar spine motion. Spine 13;1;51-53.

Giorgetti M, Harris B, & Jette (1998). Reliability of clinical balance outcome measures in the elderly. Physiotherapy Research International 3;4;274-283

Gladman J, Harwood D, Barer D (1992). Predicting the outcome of acute stroke; Prospective evaluation of five multivariate models and comparison with single methods. Journal of Neurology, Neurosurgery and Psychiatry 55;347-351.

Gladman J, Lincoln N & Barer D (1993). A randomised controlled trial of domiciliary and hospital based rehabilitation for stroke patients after discharge from hospital. Journal of Neurology, Neurosurgery and Psychiatry 56;960-6.

Gladman J, Lincoln N, & Adams SA (1993). Use of the extended ADL scale with stroke patients. Age & Ageing 22;419-424.

Gladman J, Lincoln N (1994). Follow-up of a controlled trial of domiciliary stroke rehabilitation (Domino study). Age & Ageing 23;9-13.

Goldie P, Bach T, Evans O (1989). Force platform measures for evaluating postural control: Reliability and validity. Archives of Physical Medicine and Rehabilitation 70;510-517.

Goldie PA, Matyas TA, Spencer KI & McGinley RB (1990). Postural control in standing following stroke; Test-retest reliability of some quantitative clinical tests. Physical Therapy 70;4;234-243.

Goldie P, Matyas T & Evans O (1996). Deficit and change in gait velocity during rehabilitation after stroke. Archives of Physical Medicine and Rehabilitation 77;1074-82.

Goldie P, Matyas T, Kinsella G, Galea M, Evans O & Bach T. Prediction of gait velocity in ambulatory stroke patients during rehabilitation. Archives of Physical Medicine and Rehabilitation 1999;80;415-20.

Gowland C, Stratford P, Ward M, Moreland J, Torresin W, Van Hullenaar S, Sanford J, Barreca S, Vanspall B, & Plews N (1993). Measuring physical impairment and disability with the Chedoke-McMaster stroke assessment. Stroke 24;58-63

Granger CV, Hamilton B & Sherwin F (1986). Guide for the use of the uniform data set for medical rehabilitation. Uniform Data Set System for Medical Rehabilitation Project Office, Buffalo General Hospital, New York 14203, USA

Granger CV. Cotter AC. Hamilton BB. Fiedler RC (1993). Functional assessment scales: a study of persons after stroke. Archives of Physical Medicine & Rehabilitation 74;2;133-8

Gregson J, Leathley M, Moore P, Sharma A, Smith T, Watkins C (1999). Reliability of the Tone Assessment Scale and the Modified Ashworth Scale as clinical tools for assessing post-stroke spasticity. Archives of Physical Medicine and Rehabilitation 80;1013-6.

Gregson J, Leathley M, Moore P, Sharma A, Smith T, Watkins C (2000). Reliability of measurements of muscle tone and muscle power in stroke patients. Age & Ageing 29;223-228.

Gresham G (1986). Stroke outcome research. Stroke 17;358-362.

Gresham G (1990). Past achievement and new directions in stroke outcome research. Stroke (II suppl);1S-3S.

Grimmer K (1996). The relationship between cervical resting posture and neck pain. Physiotherapy 82;1;45-51.

Gross M, Burns C, Chapman S, Hudson C, Curtis H, Lehmann J, & Renner J (1998). Reliability and validity of rigid lift and pelvic levelling device method in assessing functional leg length discrepancy. Journal of Orthopaedic and Sports Therapy 27;4;285-295

Guttman L (1944). A basis for scaling quantitative data. American Sociology Review 139-50.

Haas B & Whitmarsh T (1998). Inter- and intra- tester reliability of the Balance Performance Monitor in a non-patient population. Physiotherapy Research International 3;2;135-147

Haas B & Burden AM (2000). Validity of weight distribution and sway measurements of the Balance Performance Monitor. Physiotherapy Research International 5;1;19-32.

Halligan P, Marshall JC & Wade DT (1989). Visuospatial neglect: underlying factors and test sensitivity. Lancet 2;908-10.

Hamilton B & Granger C (1994) . Disability outcomes following inpatient rehabilitation for stroke. Physical Therapy 74;5:494-503.

Hamilton BB. Laughlin JA. Fiedler RC. Granger C (1994) . Inter-rater reliability of the 7-level functional independence measure (FIM). Scandinavian Journal of Rehabilitation Medicine 26;3:115-9.

Hanger HC, Walker G, Paterson LA, McBride S, Sainsbury R (1998) . What do patients and their carers want to know about stroke? A two-year follow-up study. Clinical Rehabilitation 12;1;45-52

Harburn K, Hill K, Kramer J, Noh S, Vandervoot A, Teasell R (1995). Clinical applicability and test-retest reliability of an external perturbation test of balance in stroke patients. Archives of Physical Medicine and Rehabilitation 76;317-23.

Hart D & Rose S (1986). Reliability of non-invasive method of measuring lumbar curve. Journal of Orthopaedic and Sports Therapy 8;180-184.

Hesse S Bertelt C, Jahnke M (1995). Treadmill training with partial body-weight support compared with physiotherapy in on-ambulant hemiparetic patients. Stroke 26;976-81.

Hill KD, Bernhardt J, McGann AM, Maltese D & Berkovits D (1996). A new test of dynamic standing balance for stroke patients; Reliability, validity and comparison with healthy elderly. Physiotherapy Canada 48;4257-62.

Hill K, Ellis P, Bernhardt J, Maggs P & Hull S (1997). Balance and mobility outcomes for stroke patients; A comprehensive audit. Australian Physiotherapy 43;3;173-180.

Hinman M (1997). Validity and reliability of measures obtained from the Balance Performance Monitor during quiet standing. Physiotherapy 83;11;579-581.

Holden MK, Gill KM, Magliozzi MR, Nathan J, Piehl-Baker L (1984). Clinical gait assessment in the neurologically impaired: reliability and meaningfulness. Physical Therapy 64;35-40

Holden M, Gill K, Magliozzi M, Nathan J, Piehl-Baker L (1984). Clinical gait assessment in the neurologically impaired: Reliability and meaningfulness. Physical Therapy 64;1;35-40.

Holmqvist L, von Koch L, Kostulas V (1998). A randomised controlled trial of rehabilitation at home after stroke in Southwest Stockholm. Stroke 29;591-7.

Horgan N & Finn A (1997). Motor recovery following stroke: A basis for evaluation Disability and Rehabilitation 19;2;64-70.

Hsieh C, Hsueh I, Mao H (2000). Validity and responsiveness of the Rivermead Mobility Index in stroke patients. Scandinavian Journal of Rehabilitation Medicine 32;140-42.

Hui E, Lum CM, Woo J (1995). Outcomes of elderly stroke patients; day hospital versus conventional medical management. Stroke 26;1616-9.

Hughes M, Duncan P, Rose D, Chandler J, Studenski S (1996). The relationship of postural sway to sensori-motor function, functional performance, and disability in the elderly. Archives of Physical Medicine and Rehabilitation 77;567-572

Indredavik B, Slordahl SA, Bakke F (1997). Stroke Unit treatment; long term effects. Stroke 28;1861-6.

Johnstone M (1987). Restoration of Motor Function in the Stroke Patient; A Physiotherapist's Approach. Churchill Livingstone, Edinburgh.

Jongbloed L, Stacey S, Brighton C (1989). Stroke rehabilitation; sensori-motor integrative treatment versus functional treatment. American Journal of Occupational Therapy 43;391-7.

Jorgensen H, Nayayama H, Raaschou H & Olsen T (1995). Recovery of walking functions in stroke patients; The Copenhagen Stroke Study. Archives of Physical Medicine and Rehabilitation 76;27-32.

Kamen G, Patten C, Du CD & Sison S (1998). An accelerometry-based system for the assessment of balance and postural sway. Gerontology 44;1;40-5.

Kalra L, Dale P & Crome P (1993). Improving stroke rehabilitation; a controlled study. Stroke 24;1462-1467.

Kalra L (1994). The influence of stroke unit rehabilitation on functional recovery from stroke. Stroke 25;821-5.

Keeley J, Mayer T, Cox R, Gatchel R, Smith J & Mooney V (1986). Quantification of lumbar function. Part 5. Reliability of range of motion measures in the sagittal plane and an in-vivo torso rotation measurement technique. Spine 11;1;31-35.

Keenan M, Perry J Jordan C (1984). Factors affecting balance and ambulation following stroke. Clinical Orthopaedics 182;165-71.

Kitzinger J (1995). Introducing Focus Groups. British Medical Journal 311;299-312.

Klaber-Moffat J, Hughes I & Griffiths P (1989). Measurement of cervical spine movements using a simple inclinometer. Physiotherapy 75;6;309-12.

Kwakkel G (1998). Dynamics in functional recovery after stroke. PhD thesis, University of Amsterdam.

Kwakkel G, Wagenaar R, Kollen B & Lankhorst (1996). Predicting disability post-stroke - A critical review of the literature. Age & Ageing 25;479-489.

Kwakkel G, Wagenaar R, Koelman TW, Lankhorst GJ, Koestier JC (1997). Effects of intensity of rehabilitation after stroke: A research synthesis. Stroke 28;1550-1556.

Kwakkel G, Wagenaar R, JW, Koelman TW, Lankhorst GJ (1999). Intensity of arm and leg training after primary middle-cerebral artery stroke; A randomised controlled trial. Lancet 354;191-6.

Langhammer B (2001). The Bobath Concept (Letters to the editor). Physiotherapy 15;l;11-13.

Langhammer B & Stranghelle J (2000). Bobath or Motor Relearning Programme? A comparison of two different approaches of physiotherapy in stroke rehabilitation; A randomised controlled study. Clinical Rehabilitation 14;361-369.

Langhorne P, Wagenaar R, Partridge C (1996). Physiotherapy after stroke: More is better? Physiotherapy Research International 1;75-88.

Le Clair K & Riach C (1996). Postural stability measures: what to measure and for how long. Clinical Biomechanics 11;3;176-178.

Lee W, Deming L & Sahgal V (1988). Quantitative and clinical measures of static standing balance in hemiparetic and normal subjects. Physical Therapy 68;6;970-977

Legh-Smith J, Wade DT, Langton-Hewer R (1986). Services for stroke patients one year after stroke. Journal of Epidemiology and Community Health 40;161-165.

Lennon S (1996). The Bobath concept; A critical review of the theoretical assumptions that guide physiotherapy practice in stroke rehabilitation. Physical Therapy Reviews 1;35-45.

Lennon S & Ashburn A (2000). The Bobath Concept in stroke rehabilitation; a focus group study of experienced physiotherapists' perspective. Disability and Rehabilitation 22;15;665-674.

Lennon S & Johnson L (2000). The modified Rivermead Mobility Index; validity and reliability. Disability and Rehabilitation 22;18;833-839

Levine D, Whittle D, Beach J & Ollard P (1996). Test -retest reliability of the Chattx Balance System in patients with hemiplegia. Journal of Rehabilitation Research & Development 33;1;36-44.

Lewis C (1993). Balance, gait tests proves simple yet useful. Physical Therapy Bulletins February 10.

Linacre JM. Heinemann AW. Wright BD. Granger CV. Hamilton BB (1994). The structure and stability of the Functional Independence Measure. Archives of Physical Medicine & Rehabilitation. 75;2:127-32.

Lincoln N & Gladman J (1992). The Extended Activities of Daily Living scale: A further validation. Disability & Rehabilitation 14;1;41-43

Lincoln N & Leadbitter D (1979). Assessment of motor function in stroke patients. Physiotherapy 65;2;48-51

Lindley R, Warlow C, Wardlaw J, Dennis M, Slattery J, Sandercock P (1993). Inter-observer reliability of a clinical classification of acute cerebral infarction. Stroke 24;1801-1804.

Linmark B & Hamrin E (1988). Evaluation of functional capacity after stroke as a basis for active intervention. Presentation of a modified chart for motor capacity assessment and its reliability. Scandinavian Journal of Rehabilitation 20;103-109.

Liston R & Brouwer B (1996). Reliability and validity of measures obtained from stroke patients using the Balance Master. Archives of Physical Medicine and Rehabilitation 77;425-430

Loewen S & Anderson B (1988). Reliability of the modified Motor Assessment Scale and the Barthel Index. Physical Therapy 68;7;1077-1081

Loewen S & Anderson B (1990). Predictors of stroke outcome using objective measurement scales. Stroke 22;1;78-81.

Lofgren B, Nyberg L, Osterlind P, Gustafson Y (1998). In-patient rehabilitation after stroke; Outcome and factors associated with improvement. Disability and Rehabilitation 20;2;55-61.

Logigan M, Samuels M, Falconer J (1983). Clinical exercise trial for stroke patients. Archives of Physical Medicine and Rehabilitation 64;364-7.

Lord J & Hall K (1986). Neuromuscular re-education versus traditional programs for stroke rehabilitation. Archives of Physical Medicine and Rehabilitation 67;88-91.

Lord SE, Halligan PW. & Wade DT (1998). Visual gait analysis: The development of a clinical assessment and scale. Clinical Rehabilitation 12;107-119.

Lord SR & Ward JA (1994). Age-associated differences in sensori-motor function and balance in community dwelling women. Age & Ageing 23;452-460.

Love S, Gringmuth R, Kazemi M, Cornacchia P & Schmolke M (1998). Inter-examiner and intra-examiner reliability of cervical passive range of motion using CROM and Cybex 320 EDI. Journal of Canadian Chiropractic Association 42;4;222-8.

Lynch M & Grisigno V (1991). Stroke and Head Injuries: A Guide for Patients, Families, Friends and Carers. J Murray, London

Lynch SM, Leahy P & Barker SP (1998). Reliability of measurements obtained with a modified functional reach test in subjects with spinal cord injury. Physical Therapy 78;2;128-133

Malouin F, Pichard L, Bonneau C, Durand A & Corriveau D (1994). Evaluating motor recovery early after stroke; comparison of the Fugel Meyer Assessment and Motor Assessment Scale. Archives of Physical Medicine and Rehabilitation 75;1206-1212.

Martin J, White A, Meltzer H (1989) Office of population censuses and services: Disabled adults, services, transport and employment (report 4, disability in Great

Matyas TA & Ottenbacher KJ (1993). Confounds of insensitivity and blind luck; statistical conclusion validity in stroke rehabilitation. Archives of Physical Medicine & Rehabilitation 74;559-565.

Matyas N & Pope C (1995). Observational methods in health care settings. British Medical Journal 311;182-4.

Mayer T, Tencer A, Kristoferson S, Mooney V (1984). Use of non-invasive techniques for quantification of spinal range of motion in normal subjects and chronic low back pain patients. Spine 9;6;588-595.

Mayer T, Kondraske G, Beals S, Gatchel R (1997). Spinal range of motion; Accuracy and sources of error with inclinometric measurement. Spine 122;171976-1984.

Mayer S. Chen I, Lavender S, Trafimow J, Andersson (1995). Variance in the measurement of sagittal lumbar spine range of motion among examiners, subjects and instruments. Spine 13;1489-1493.

Mayo N, Korner-Bitensky N, Becker R (1991). Recovery time of independent function post-stroke. American Journal of Physical Medicine & Rehabilitation 70;1;5-12.

Mayston M (2000a). Motor learning now needs meaningful goals *(Letters to the editor)* Physiotherapy 86;9;492.

Mayston M (2000b). Compensating for CNS dysfunction *(Letters to the editor)* Physiotherapy 86;11;612

McLean I, Gillan M, Ross J, Aspden R, Porter R (1996). A comparison of methods for measuring trunk list. A simple plumbline is the best. Spine 21;14;1667-1670.

McKensie C & Charlson M (1986). Standards for use of ordinal scales in clinical trials. British Medical Journal 292;40-3.

McKnight C & Rockwood K (1995). A hierarchical assessment of balance and mobility. Age & Ageing 24;126-130.

Mellin G (1986). Measurement of thoraco-lumbar posture and mobility using a Myrin inclinometer. Spine 11;7;759-762.

Mellin G, Kiisi R, Weckstrom A (1991). Effects of subject position on measurements of flexion, extension and lateral flexion of the spine. Spine 16;9;1108-1110.

Menzel H (1953). A new coefficient of scalogram analysis. Public Opinion Quarterly 17;268-80.

Moll J, Liyange S & Wright V (1972). An objective clinical method to measure lateral spinal flexion. Rheumatology and Physical Medicine 11;225-239.

Moll J, Liyange S & Wright V (1972). An objective clinical method to measure spinal extension. Rheumatology and Physical Medicine11;293-312

Morgan DL (1997). Focus groups As Qualitative Research (2nd ed). Sage Publishing, London.

Morgan P (1994). The relationship between sitting balance and mobility outcome. Australian Journal of Physiotherapy 40;2;91-96.

Nattrass C, Nitske J, Disler P, Chou M & Ooi K (1999). Lumbar spine range of motion as a measure of physical and functional; impairment; an investigation of validity. Clinical Rehabilitation 13;211-218.

Nelson DL, Konosky K, Fleharty K (1996). The effects of an occupationally embedded exercise on bilaterallly assisted supination in people with hemiplegia. American Journal of Occupational Therapy 50;639-46.

Newton J & Waddell G (1991). Reliability and validity of clinical measurement of the lumbar spine in patients with chronic low back pain. Physiotherapy 12;769-800

Niam S. Cheung W. Sullivan PE. Kent S. Gu X (1999). Balance and physical impairment after stroke. Archives of Physical Medicine and Rehabilitation

80;10;1227-33.

Nichols D, Glenn T, Hutchinson K (1995). Changes in the mean centre of balance during balance testing in young adults. Physical Therapy 75;699-706.
Nichols D, Miller L, Colby L & Pease W (1996). Sitting balance: Its relation to function in individuals with hemiparesis. Archives of Physical Medicine and Rehabilitation 77;865-7

Nieuwboer A Feys H, De Weerdt W, Nuyens G, DeCorte E (1995). Developing a clinical tool to measure sitting balance after stroke: A reliability study. Physiotherapy 81;8;439-445.

Nilsson L & Nordholm L (1992). Physical therapy in stroke rehabilitation; Bases for Swedish physiotherapists' choice of treatment. Physiotherapy Theory and Practice 8;49-55

Nilsson L, Carlsson J, Grimby G & Nordholm L (1998). Assessment of walking, balance and sensori-motor performance of hemiparetic patients in the acute stage after stroke. Physiotherapy Theory and Practice 14;149-157.

Nitske J, Nattrass C, Disler P, Chou M & Ooi K (1999). Reliability of the American Medical Association guides model of measuring spinal range of motion. Spine 24;3;262-268.

Nitz J & Gage A (1995). Post stroke recovery of balanced sitting and ambulation ability. Australian Journal of Physiotherapy 41;4;263-267

Norusis M (1997) SPSS 6.1, Guide to data analysis. Simon & Schuster. New Jersey.

Nouri F & Lincoln N (1987). An extended activities of daily living scale for stroke patients. Clinical Rehabilitation 1;301-5.

Office of Health Economics (1988) Stroke A series of papers on current health problems). London: Office of Health Economics

Pai Y, Rogers M, Hedman L, Hanke T (1994). Alterations in weight-transfer capabilities in adults with hemiparesis. Physical Therapy 74;7;647-659.

Pandyan A, Johnson G, Price C, Curless R, Barnes M & Rodgers H (1999). A review of the properties and limitations of the Ashworth scale as measurements of spasticity. Clinical Rehabilitation 13;373-383.

Panturin E (2001). The Bobath Concept (Letter to the editor) Clinical Rehabilitation 15;111-13.

Paolucci S, Antonucci G, Grasso M, Pizzamaglio L (2001). The role of unilateral spatial neglect in rehabilitation of right brain damaged ischaemic stroke patients; a matched comparison. Archives of Physical Medicine and Rehabilitation;82;6;743-9.

Paolucci S, Grasso M, Antonucci G, Bragoni M, Troisi E, Morelli D, De Angelis D & Rizzi F (2001). Mobility status after inpatient stroke rehabilitation: 1-year follow up and prognostic factors. Archives of Physical Medicine and Rehabilitation 82;2-8.

Partridge C, Johnston M & Edwards S (1987). Recovery from physical disability after stroke; Normal patterns as a basis for evaluation. Lancet i;373-75

Partridge C & Edwards S (1988). Recovery curves as a basis for evaluation. Physiotherapy 74;3;141-143

Partridge C, Morris L & Edwards S (1993). Recovery from physical disability after stroke; profiles for different levels of starting severity. Clinical Rehabilitation 7;210-217

Partridge C (1994) Evaluation of physiotherapy for people with stroke. Kings Fund, London

Partridge C, Cornall C, Lynch M, Greenwood R (1997). Physical Therapies. in Neurological Rehabilitation (eds Greenwood R, Barnes M, McMillan T, Ward C) Mosby Hall, London.

Patel M, Potter J, Perez I, Kalra L(1998). The process of rehabilitation and discharge planning in stroke; a comparison between stroke units. Stroke 29;2484-7.

Podsiadlo D & Richardson S (1991). The timed up and go. A test of basic functional mobility for frail elderly persons. Journal of the American Geriatric Society 39;142

Pollock A, Durward B, Rowe P (1998). Symmetry of weight distribution during sitting in subjects with recently acquired stroke. Proceeding of Physiotherapy Research Society in Physiotherapy 84;8;402.

Pomeroy V, Dean D, Sykes L, Faragher B, Yates M, Tyrrel P, Moss S, Tallis R (2000). The unreliability of clinical measures of muscle tone: implications for stroke therapy. Age & Ageing 29;229-233.

Pomeroy V & Tallis R (2000). Physical therapy to improve movement performance and functional ability post-stroke. Part 1. Existing evidence. Reviews in Clinical Gerontology 10;261-290

Pomeroy V, Niven DS, Barrow S, Faragher E, Tallis R (2001). Unpacking the black box of nursing and therapy practice for post-stroke shoulder pain: a precursor to evaluation. Clinical Rehabilitation 15;67-83.

Poole J & Whitney S (1988). Motor Assessment Scale for stroke patients: Concurrent validity and inter-rater reliability. Archives of Physical Medicine and Rehabilitation 69;195-197

Prescott R, Garraway W, Akhtar A (1982). Prediciting functional outcome following acute stroke using a standard clinical examination. Stroke13;5;641-647.

Rappaport J & Eerd MJ (1989). Impact of physical therapy weekend coverage on length of stay in an acute care community hospital. Physical Therapy 69;32-7.

Rankin G & Stokes M (1998). Reliability of assessment tools in rehabilitation; an illustration of appropriate statistical analyses. Clinical Rehabilitation 12;187-199.

Reding M & Potes E (1988). Rehabilitation outcome following initial unilateral hemispheric stroke; Life table analysis approach. Stroke 19;1354-1358.

Reynolds P (1975). Measurement of spinal mobility: A reliability comparison of three methods. Rheumatology and Rehabilitation 14;180-185.

Richards CL, Malouin F, Wood-Dauphinee S (1993). Task-specific therapy for optimisation of gait recovery in acute stroke patients. Archives of Physical Medicine and Rehabilitation 74;612-20.

Rodgers H, Soutter J, Kaiser W (1997). Early supported discharge following acute stroke; pilot study results. Clinical Rehabilitation 11;280-7.

Rondinelli R, Murphy J, Esler A, Marciano T & Cholmakjian C. Estimation of normal lumbar flexion with surface inclinometry; A comparison of three methods. American Journal of Physical Medicine & Rehabilitation 1992;71;4;219-24.

Ronning OM & Guldvog B (1998a). Stroke unit versus general medical wards, II: neurological deficits and activities of daily living. A quasi-randomised controlled trial. Stroke 29;586-90.

Ronning OM & GuldvogB (1998b). Outcome of subacute stroke rehabilitation; A randomised controlled study. Stroke 29;779-84.

Rossier P & Wade DT (2001). Validity and reliability of 4 mobility measures in-patients presenting with neurological impairment. Archives of Physical Medicine and Rehabilitation 82;9-13.

Rothstein J. (1993). Truth and Theory. Ch 1 in Research in Physical Therapy. (ed) CE Bork. JB Lippincott Ltd.

Royal College of Physicians (2000) National Clinical Guidelines for Stroke. Royal College of Physicians, London

Rudd AG, Wolfe CD, Tilling K, Beech R (1997). Randomised controlled trial to evaluate early discharge scheme for people with stroke. British Medical Journal 315;1039-44.

Sackley C (1990). The relationship between weight-bearing asymmetry after stroke, motor function and activities of daily living. Physiotherapy Theory & practice 6;179-185.

Sackley C & Lincoln N (1990). The verbal administration of the gross function section of the Rivermead motor assessment. Clinical Rehabilitation 4;301-303.

Sackley C & Lincoln N (1991). Weight distribution and postural sway in healthy adults. Clinical Rehabilitation 5;181-186.

Sackley C & Lincoln N (1996). Physiotherapy treatment for stroke patients; A survey of current practice. Physiotherapy Theory and Practice 12;87-96.1996

Sanchez-Blanco I, Ochoa-Sangrador C, Lopez-Munain L, Izquierdo-Sanchez, Fermoso-Garcia (1999). Predictive Model of functional independence in stroke patients admitted to a rehabilitation unit. Clinical Rehabilitation 13;464-475.

Sandford J, Moreland J, Swanson P, Gowland C (1993). Reliability of the Fugel-Meyer assessment for testing motor performance in patients following stroke. Physical Therapy 73;7;447-454

Sandin K & Smith B (1990). The measure of balance in sitting in stroke rehabilitation prognosis. Stroke 21;82-86.

Saur P, Ensink F, Frese K, Seegar D & Hildebrandt J (1996). Lumbar range of motion; Reliability and validity of the inclinometer technique in the clinical measurement of trunk flexibility. Spine 21;11;1332-1338.

Sawner K & LeVigne J (1992) Brunnstrom's Movement Therapy in Hemiplegia: A neurophysiological approach. (2nd ed). J. JB Lippincott Co. Philadelphia

Shelley P (2000). Clarification needed on handling. *(Letters to the editor)* Physiotherapy 86;9;494.

Shepherd K (1991). Theory; criteria, importance and impact. Contemporary management of motor control problems. Proceedings of the II Step conference 5-10. APTA, Virginia.

Shrout P & Fleiss J (1979). Intra-class Correlations; Uses in assessing rater reliability. Psychological Bulletin 86;420-428.

Shumway-Cook A & Woolacott M (1995). Motor Control Theory & Application. Williams & Wilkins, Baltimore USA.

Shumway-Cooke A & Woolacott M (2000) Motor Control Theory and Practical Application (2nd ed). Lippincott Williams & Wilkins, Philadelphia

Shumway-Cook A & Horak F (1986). Assessing the influence of sensory interaction on balance; Suggestions from the field. Physical Therapy 66;1548-1550.

Shumway-Cook A, Brauer S & Woolacott M (2000). Predicting the probability of falls in community-dwelling older adults using the Timed Up & Go Test. Physical Therapy 80;9;869-903.

Simondson J, Goldie P, Brock K, Nosworthy J (1996). The Mobility Scale for Acute Stroke Patients; Intra and inter-rater reliability. Clinical Rehabilitation 10;295-300.

Skilbeck C, Wade DT, Langton Hewer R, Wood V (1983). Recovery after stroke. Journal of Neurology, Neurosurgery and Psychiatry 46;5-8.

Smith DS, Goldenberg E, Ashburn A, Kinsella G, Shiekh K, Brennan P. Remedial therapy after stroke; A randomised controlled trial. British Medical Journal (1981) 282;517-20.

Smith J, Brotheridge S, & Young J (2001). Patterns of hemiparesis recovery in lacunar and partial anterior circulation infarct stroke syndromes. Clinical Rehabilitation 15;59-66.

Smith M & Baer G (1999). Achievement of simple mobility milestones after stroke. Archives of Physical Medicine and Rehabilitation 80;442-7.

Smith R (1994). Validation and reliability of the Elderly Mobility Scale. Physiotherapy 80;11;744-747

Stevenson T & Garland J (1996). Standing balance during internally produced perturbations in subjects with hemiplegia: validation of the balance scale. Archives of Physical Medicine & Rehabilitation 77;656-662

Streiner D & Norman G (1995). Health Measurement Scales; A Practical Guide to Their Development and Use (2nd ed). Oxford Medical Publications.

Stroke Unit Trialists Collaboration (1998). Organised inpatient (stroke unit) care for stroke (Cochrane Review) in The Cochrane Library. Issue 3 1999.

Tabachnick B & Fidell L (1996). Using multivariate statistics. 3rd edition. Harper Collins. New York.

Taylor B, Ellis E, Haran D (1995). The reliability of measurement of postural alignment to assess muscle tone change. Physiotherapy 81;8;485-490.

Thorbahn L & Newton R (1996). Use of the Berg Balance Scale to predict falls in the elderly. Physical Therapy 76;6;576-583.

Tinetti M (1986). Performance orientated assessment of mobility problems in elderly patients. Journal of American Geriatric Society 34;119-126.

Tinetti ME, Speechley M & Ginter SE (1988). Risk factors for falls in among elderly persons living in the community. New England Journal of Medicine 319;1701-1707

Tucci S, Hicks J, Gross E, Campbell W & Danoff J (1986). Archives of Physical Medicine & Rehabilitation 67;225-230.

Turner-Stokes L (1999). The effectiveness of rehabilitation; a critical review of the evidence, introduction. Clinical Rehabilitation 13;(suppl 1);3-6.

Tyson SF & Turner GF (1999). The process of stroke rehabilitation; What happens and why? Clinical Rehabilitation 13;322-332

Tyson SF & Turner GF (2000). Discharge and follow-up for people with stroke; What happens and why. Clinical Rehabilitation 14;381-392.

Wade DT (1992). Measurement in Neurological Rehabilitation. Oxford University Press, Oxford.

Wade DT (1994). Stroke (acute cerebro-vascular disease). In Stevens A, Raftery J (eds) Health care needs assessment: the epidemiology based needs assessment reviews. Oxford: Radcliffe Medical Press

Wade DT (1998a). Evidence relating to goal planning in rehabilitation. Clinical Rehabilitation 12;273-5.

Wade DT (1998b). A framework for considering rehabilitation interventions. Clinical Rehabilitation 12;363-368.
Wade DT (1999). Evidence relating to goal planning in stroke rehabilitation; How? Topics in Stroke Rehabilitation 6;16-36.

Wade DT, Collen FM, Robb GF, Warlow CP (1992). Physiotherapy intervention late after stroke and mobility. British Medical Journal 304;609-13.

Wade DT & Collin C (1988). The Barthel ADL Index; A standard measure of physical activity. International Disability Studies 10;64-67.

Wade DT & Langton-Hewer R et al (1987). Motor loss and swallowing difficulty after stroke: Frequency, recovery and prognosis. Acta Neurologica Scandinavia 76;50-4.

Wade DT & Langton-Hewer R (1987). Functional abilities after stroke: Measurement, natural history and prognosis. Journal of Neurology, Neurosurgery and Psychiatry 50;177-182.

Wade DT, Langton-Hewer R, Skilbeck CE, David RM (1985). Stroke: A critical approach to diagnosis, treatment and management. Chapman & Hall Medical. London

Wade DT, Skilbeck C, Langton-Hewer R, Woods V (1984). Therapy after stroke: amounts, determinants and effects. International Rehabilitation Medicine 6;105-10.

Wade DT, Wood V, Langton Hewer R (1985). Recovery after stroke - The first three months. Journal of Neurology, Neurosurgery and Psychiatry 48;7-13.

Wade D, Wood V, Heller A, Maggs J, Langton Hewer R (1987). Walking after stroke; Measurement and recovery in the first three months. Scandinavian Journal of Rehabilitation Medicine 19;25-30.

Wagenaar RC, Meijer O, van Wieringen PCW (1990). The functional recovery of stroke; a comparison between neuro-developmental treatment and the Brunnstrom method. Scandinavian Journal of Rehabilitation Medicine 22;1-8.et al

Wagenaar R & Meijer O (1991a). Effects of Rehabilitation (1). Journal of Rehabilitation Science 4;61-73.

Wagenaar R & Meijer O (1991b). Effects of Rehabilitation (2). Journal of Rehabilitation Science 4;97-109

Walker M, Rothstein J, Finucane S, Lamb R (1987). Relationships between lumbar lordosis, pelvic tilt and abdominal performance. Physical Therapy 67;512-516.

Wee J, Bagg S, Palepu A (1999). The Berg Balance Scale as a predictor of length of stay and discharge destination in an acute stroke rehabilitation setting. Archives of Physical Medicine and Rehabilitation 80;448-52.

Weiner DK, Duncan PM, Chandler J, Studenski S (1992). Functional Reach: A marker of physical frailty. Journal of American Geriatric Society 40;203-207.

Weiner D, Bongiorni D Studenski S, Duncan P, Kochersberger G (1993). Does functional reach improve with rehabilitation? Archives of Physical Medicine and Rehabilitation 74;796-800.

Wilson B, Cockburn J, & Halligan P (1987). Development of a behavioural test of visuospatial neglect. Thames Valley Test Co.

Wilson B, Cockburn J & Halligan P (1987). Development of a behavioural test of visuo-spatial neglect. Archives of Physical Medicine and Rehabilitation 68;98-102

Wilson B, Cockburn J & Halligan P (1990). A short screening test for visual neglect in stroke patients. International Disability Studies 12;95-99.

Wing A, Barton J, Turton A & Howick I (1992). Regulation of lateral position of body centre of mass in standing balance. Physiotherapy Theory & practice 8;131-135

Wing A, Goodrich S, Virji-Babul N, Jenner J, Clapp S (1993). Balance evaluation in hemiparetic stroke patients using lateral forces applied to the hip. Archives of Physical Medicine and Rehabilitation 74;292-299.

Winstein C, Gardner E, McNeal D, Barto P, Nicholson D (1989). Standing balance training: Effect on balance and locomotion in hemiparetic patients. Archives of Physical Medicine and Rehabilitation 70;755-762

Winward C, Halligan P & Wade DT (2000). Rivermead Assessment of Somatosensory Performance. Thames Valley Testing Co. Bury St Edmunds.

Wolfson L, Whipple R, Amerman P, Kleinberg A (1986). Stressing the postural response - quantitative method for testing balance. Journal of American Geriatrics Society.;34;845-50.

Wood-Dauphinee S, Williams J & Shapiro S (1990). Examining outcome measures in a clinical study of stroke. Stroke 21;731-739

Wood-Dauphinee S, Berg K, Bravo G & Williams J (1996). The Balance Scale: Responsiveness to clinically meaningful changes. Canadian Journal of Rehabilitation 10;1;35-50.

Wright B (1971). A simple mechanical ataxia-meter. Journal of Physiology;218;27-28P

Young J & Forster A (1992). The Bradford community stroke trial: Results at six months. British Medical Journal 304;1085-9.

Young J & Forster A (1993). Day hospital and home physiotherapy for stroke patients; a comparative cost-effectiveness study. Journal of the Royal College of Physicians, London 27;252-8.

Youdas J, Carey J & Garrett T (1991). Reliability of measurements of cervical spine range of motion - comparison of three methods. Physical Therapy 71;2;98-104.

Youdas J, Susman V & Garrett T (1995). Reliability of measurements of the lumbar spine sagittal mobility obtained with the flexible curve. Journal of Orthopaedic and Sports Therapy 21;1;13-20.

Youdas J, Garrett T, Harmsen S, Susman V & Carey J (1996). Lumbar lordosis and pelvic inclination of symptomatic adults. Physical Therapy 76;10;1066-1081.

Appendix I
Data from the focus groups showing categorisation

Listed below are the data from the focus groups. These were taken from the flip charts, transcribed verbatim on to a word processing package and grouped in to the categories. The categories are: 1. Purpose of assessment, 2. Process and clinical reasoning 3. Alignment of body segments the physiotherapists observed 4. Aspects of weight distribution the physiotherapists observed 5. Aspects of muscle activity the physiotherapists tested 6. Other issues not included in the other categories

1. Purpose of Assessment

What is primary problem?
What are compensations?
What are the reasons for abnormalities
What looks wrong?
What muscle work occurring? Why? What compensations are used? Can tell by see and feel?
Why is the posture flexed? - ?leaning ?fixed
Why flexed posture - ?leaning ?fixed ?weak trunk
Why is the side flexed ?collapsed, ?pulled down
Why is the trunk rotated ?pulling ?collapse
Why flexed on weak side - ?latisimus dorsi, ?pulled down ?collapse
Is it an effort to maintain the posture or move from it? Is it comfortable? Difficult?
Can he function?
Are his hips active? Is it increased to stabilise? What is happening to the pelvis?
Is it free or compensation for stability?
Is the head, legs, arms fixed to maintain position or free to move?
Can he do something else?
Is he using hands for support?
Can he move from posture?
Can he move around base of support?
How does he react to touch and handling
Can he let go of contraction?
Can he relax to command?
Can you alter posture using different key points by handling?
Can he grade, co-ordinate movement?
What does it feel like? Feel for pull in to flexion

2. Process and Clinical reasoning

Comparisons to aid diagnosis

Position relative to expectation
Compare with normal
Compare alignment of body segments relative to each other
Compare with his normal
Alignment - Compare with normal, and alignment of segments relative to each other

Questions the physiotherapists ask themselves

Can he correct position with verbal prompt?
What is primary problem? What are compensations?
What are the reasons for abnormalities?
What looks wrong?

What muscle work occurring? Why? What compensations are used? Can tell by see and feel?
Where does extension come from?
Dynamic stability around joint – muscle activity should be adaptable, co-ordinated. harmonious = co-contraction. Not locked in closed packed position.

Why is the posture flexed? - ?leaning ?fixed
Why flexed posture - ?leaning ?fixed ?weak trunk
Why is he side flexed ?collapsed, ?pulled down
Why is the trunk rotated ?pulling ?collapse
Why flexed on weak side - ?latisimuss dorsi, ?pulled down ?collapse

How much activity is there in the foot?
Can he take weight through the foot in standing?
Is the foot in contact with the floor? If not why not?
Is the foot position secondary to hip/knee position?
Is the knee flexed? Is it active or secondary to hip/pelvic position?
Are the knees, feet, toes stable?
What is causing the hip retraction?

Feet on floor – pushing from S. side, clawing of toes
Is he weight bearing through foot –flopping or fixing?
Can he adapt to change in position of the foot – with pronation/ supination

Is the position of the legs (rotation, abduction/adduction, knee flexion) fixed or relaxed?
Is he able to accept his foot on the floor?
Is his foot on the floor? How? - pushing from sound side, clawing of toes?
Is the weight bearing foot flopping or fixing?
? Able to adapt to change in position – with pronation/ supination
Is there muscle shortening?
Is mobility limited by stiffness, shortening?
Mobility of head and limbs – free for spontaneous movement or hanging on?
Is he using hands for support?
Are the arms stuck or free?
Is he using arms for support?
Are the arms free or fixed in compensation to maintain stability of trunk?
Why isn't he weight bearing on Weak side? Weight on sound side → ↑ activity on Sound Side
Weight transfer through bottom. Is he leaning, pushing, or accepting Base of Support?
Weight transfer through feet – Is weight going through floor? Is he accepting Base of support?
Is there soft tissue shortening? Is it affecting the posture?
Is there muscle shortening?
Is it effortless?
What muscle activity to transfer weight? Is he hitched with good side?
Does he use effort to maintain posture or move from it? Is it comfortable? Difficult?
Is he able to function?
What is the activity of hip muscles? Is it increased to stabilise? Is it influencing the influencing pelvis?
Is the activity free or compensation for stability?

Is the position of leg free or fixed? Is it a compensation to stabilise/ maintain posture?
Is he using abduction/adduction to stabilise?
Are the toes curling?
Is he fixing with head, legs, or arms to maintain position or is he free to move or do something else?
Is it an effort to maintain the posture or move from it? Is it comfortable? Difficult?
Can he function?
Are his hips active? Is it increased to stabilise? What is happening to the pelvis?
Is it free or compensation for stability?
Is the head, legs, arms fixed to maintain position or free to move?
Can he do something else?
Is he using hands for support?
Can he move from posture?
Can he move around Base of support?
How does he react to touch and handling
Can he let go of contraction?
Can he relax to command?
Can you alter posture using different key points by handling?
Can he grade, co-ordinate movement?
What does it feel like? Feel for pull in to flexion
Is he fixing eyes to compensate?
Is he aware of WEIGHT BEARING?
Is his posture long standing?

Movements the physiotherapists observe and test

Movement patterns – effect on alignment of key points
? can move from posture
? Can move around Base of support
Feel movement and resistance to movement
Bony points as markers of segments
Guide movement –
Handle to see if can alter posture using different key points. ?Can grade, co-ordinate movement
Quality of movement – smooth, effortless, rhythm
Reaction to touch and handling ?Can let go of contraction / relax to command

3. Alignment of body segments that the physiotherapists observe

Position / Alignment of Head & Neck

Head and neck posture – poking chin
Position of head: trunk
Head position
Joint range of head/ neck,
Alignment – head on shoulders,
Head rotated to sound side
Neck posture – flexion /ext
Contours of neck/ shoulder and head on shoulder
Lacks head righting
Head righting – eyes level, head on trunk position

↓ head righting, using more extension
↓ head on trunk righting reactions
↓ head righting – ↑ neck muscle activity on sound side to compensate
No head righting and flexed posture
Head righting
No head righting reactions

<u>Position / Alignment of Trunk</u>
Thoracic position – Flexion, extension, rotation
Trunk extension – lumbar and thoracic extension
Side flexion on weight bearing side and elongation on non-weight bearing side
Alignment – trunk over pelvis
Side flexed, lacks elongation on W. side
Upper on lower trunk
Alignment Shoulders on pelvis
Flexed posture, leant forwards
Side flexed on right
Lumbar spine still neutral, side flexion at lumbar/thoracic junction
Lumbar spine flexion, side flexion
Kyphosis
Not extending on weight bearing side
Head forwards - trunk flexed
Side flexed trunk
Lacks extension on weight bearing side
Kyphosis/ Lordosis with pelvic tilt
Anterior tilt/ lumbar extension ? compensates with thoracic extension and shoulder girdle retraction
Flexed ++ - passive
No elongation on sound side
Weak side flexion missing
Weak trunk – side flexed ?collapsed, ?pulled down
Lumbar spine stiff
Flexion/ extension Lumbar spine – range and activity
Lack of elongation on weight bearing side
Lack of side flexion on non-weight bearing side
Lacks active trunk extension
? Leaning forwards on left
Trunk rotation ?pulling ?collapse
Alignment of trunk on pelvis – side flexion, flex/ext, rotation
Vertical alignment of trunk/ spine on pelvis and head on trunk
Side flexion of the trunk
Alignment of key points Central key point/trunk: Shoulder girdle, Central key point /trunk : pelvis – lateral and anterior / posterior
Trunk posture – flex/ext,
Alignment and symmetry of central key point: Proximal Key point (pelvis, shoulder girdle) : Distal Key Point
Dissociation head/ trunk
Alignment (head:trunk, trunk:hip/pelvis)

Waist creases
Skin creases the same -> not shifted weight
Skin folds of waist
Skin folds – waist, abdominals, level of umbilicus, and upper limb
Skin creases
Upper trunk posture
Lead by lumbar spine
Falling to weak side, collapse →Can't maintain support/ position on W. side
Need correct ratio between trunk and pelvis
Extension shoulder girdle → extension of weight bearing side
↓ trunk on pelvis right reactions - side flexion with lateral tilt, lower limb – extension/abduction
Displacement of weight and reaction → righting reactions
Trunk righting – side flexion, flexion/extension, rotation
Flexed on weak side - ?latisimuss dorsi, ?pulled down ?collapse
Bony alignment Spine: pelvis
Side flexion/ elongation on Sound/ Weak side
Righting reactions
No trunk righting reactions

<u>Position or Alignment of Pelvis/ Hips</u>
Pelvic tilt – anterior / posterior
Lateral tilt of pelvis, up on non-weight bearing side
Neutral anterior/ posterior tilt
Lateral tilt of pelvis x 2
Pelvis still neutral
Anterior / posterior tilt of pelvis
Very little lateral pelvis tilt
Not lateral tilted pelvis
Can't sustain anterior / posterior tilt
Levels of Posterior superior iliac spines / iliac crests → Pelvic tilt –anterior / posterior /lateral
Lacks anterior tilt of pelvis
Range in pelvis –anterior/ posterior tilt, lateral tilt
No anterior tilt of pelvis or lumbar extension
Minimal lateral tilt
Lateral and anterior / posterior pelvic tilt – range and activity
No pelvic tilt
Lacks lateral tilt and weight shift
Anterior / posterior position of pelvis
Pelvis – lacks eccentric extension on weak side and side flexion on non-weight bearing side
No change in pelvic position – kept anterior and lateral tilt
Position of pelvis on hip – anterior/posterior tilt, lateral tilt, retracted, adduction/medial rotation/flexion
Head flexed and side flexed → rotation to Weak side
→ pelvis collapsed – lateral tilt, rotated forwards
→ trunk rotation and side flexed on S. side

Leg and Foot Position/ Alignment

Foot position
Hip position – ab/adduction
Position of Legs – expect abduction /medial rotation on non-weight bearing side, adduction /lateral rotation on weight-bearing side
? Feet on floor – pushing from sound side, clawing of toes
  weight bearing foot –flopping or fixing
  ? Able to adapt to change in position – with pronation/ supination
Feet on floor ? able to accept
foot flat on floor
Position of legs – rotation, ab/adduction, knee flexion, ?relaxed or fixed
Feet – contact with floor,
Foot – level of activity. ?Could take weight through it in standing
Weak knee sagged → hip and knee flexed
Ab/adduction and flex/ext of hips
? Foot contact, if not why not- ?secondary to hip/knee position
Knee flexed ?active, ?secondary to hip/pelvic position.
? Stability of knee, feet, toes
Lower limb alignment – knees, feet
Hip retraction
  ? Feet on floor – pushing from sound side, clawing of toes
  weight bearing foot –flopping or fixing
  ? Able to adapt to change in position – with pronation/ supination

Position or Alignment of Scapula and Upper Limb

Shoulder retraction
Bony alignment – spine: scapular
Joint range - shoulder, elbow
Rotation retraction
Weak scapula wasted, rotated
Sound scapula more elevated and active
Elevation shoulder on Sound side
No arm righting – clamped to side to stabilise
Wasting weak shoulder
Position of arms and hands relative to trunk
Subluxed shoulder
Position of scapulae: Spine (inferior angle and spine: spinal process)
Weak scapula winged and wasted – inc. activity of trapezius
Position and height of scapular on chest wall
Distance of upper limb from trunk
Upper limb – free not fixed, relaxed not supporting
S. arm has come away from body, less fixing in adduction
Scapular retracted – high tone in pectorals, trapezius, levator scapulae. ?shortened
UL alignment on trunk – rotation, flexion, abduction
Scapular level, position, winging
L arm in same position – stays adducted
Position of arm/ scapula relative to trunk
Position of shoulder girdle on trunk

Scapulae – height and position on chest wall
Arm relative to trunk – distance and symmetry
Weak shoulder subluxed, heavy, low tone
Position of arms related to trunk
Sound arm adducted and medially rotated – pulled into fix
Position of arms, legs, hands, feet.
Limbs relative to trunk and pelvis

Symmetry
Compare Right:Left
Symmetry x 7
Should be more asymmetrical
Mid-line – where is it?
Symmetry of tummy,
Symmetry of ribs (level, contours)

4. Aspects of weight distribution that the physiotherapists observe

Sway and movement, strategies to keep balance
Weight distribution x 3
Weight transfer and distribution on weight bearing side
Very little weight transfer
Big shift on Centre of gravity within Base of support
Little shift of Centre of gravity within Base of support
Position of Centre of gravity within Base of support – sacral sitting
Falling back and to weak side
weight bearing more on Sound side
Weight transfer to weight bearing side lacking
Not weight bearing on Weak side – weight on sound side → ↑ activity on S Side
Weight transfer
Weight distribution –laterally and anterior / posterior
Weight transfer – through bottom - ? lean ?push ?accepting Base of support
Weight transfer - through feet - ?weight going through floor ?accepting Base of support
Weight distribution in lateral weight transfers, narrow Base of support, and step standing
↓ weight transfer to weak side
Base of support – size, position of Centre of gravity within Base of support, taking weight,
Static <-> weight transfers <-> single stance

5. Aspects of muscle activity that the physiotherapists test

Tone
Low tone x3
Muscle tone - symmetry, visible tension esp. in paravertebral
Changes in tone when move
Abdominal tone,
Influence of gravity on: Muscle tone

↑or ↓ tone
↑ tone and associated reactions (pathological) on weak side – arm and leg
Palpate muscles – neck, waist, scapula, erector spinae
Muscle bulk – atrophy, hypertrophy, spasticity

Muscle Length
Palpate muscle groups – muscle activity and length
Soft tissue shortening ?effect on posture
Knee extension and TA limited – lack of active movement, soft tissue short,
? muscle shortening
Muscle length of pectorals, latissimus dorsi, abdominals, hip flexors, adductors, hamstrings, lumbar spine extensors
Mobility - ? limited by stiffness, shortening

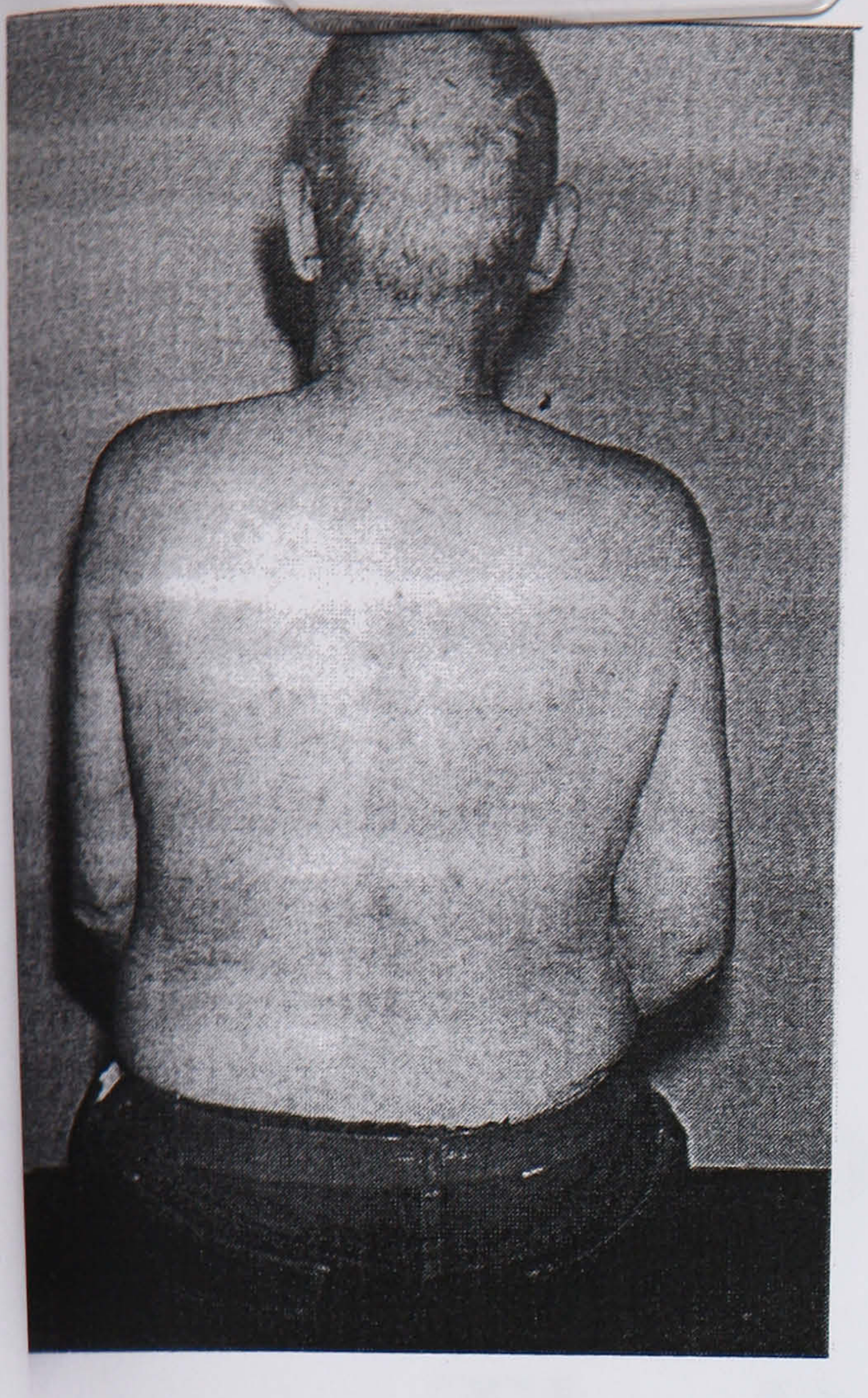
Free or Fixed? Compensatory Position or Increased Activity
↑ activity on sound side
Arms – level of activity
Degree of difficult maintaining position
? effortless
↑ activity erector spinae sound side >weak side
↑ Activity in neck
Muscle activity to maintain position
↑ activity on sound side to compensate
Muscle activity to transfer weight ?hitched with good side
Position, and muscle activity in Lower limbs
Sound leg inc. activity to stabilise
Abdominal and trunk muscle activity
? Effort to maintain posture or move from it/ ?comfortable, difficult, able to function
?Activity of hip muscles – inc. to stabilise, ?influencing pelvis
Fixing S. shoulder - -over active to compensate
Overactive on sound side to stop falling
Overactive on sound side to compensate.
Fixing and overusing Sound shoulder and pectorals to compensate
Fixing on sound side with upper limb
? Free or compensation for stability
Position of leg – free or fixed – to compensate stabilise/ maintain posture ? In ab/add
? Toes curling
? fixing with head, legs, arms to maintain position or free to move/ do something else
Free to move/use arms, head, feet or fixed to stabilise
Fixed / stability
Compensation / fixing
? Effort to maintain posture or move from it/ ?comfortable, difficult, able to function
?Activity of hip muscles – inc. to stabilise, ?influencing pelvis
? Free or compensation for stability
Position of leg – free or fixed – to compensate stabilise/ maintain posture ? In ab/add
? Toes curling
? fixing with head, legs, arms to maintain position or free to move/ do something else
? Using hands for support

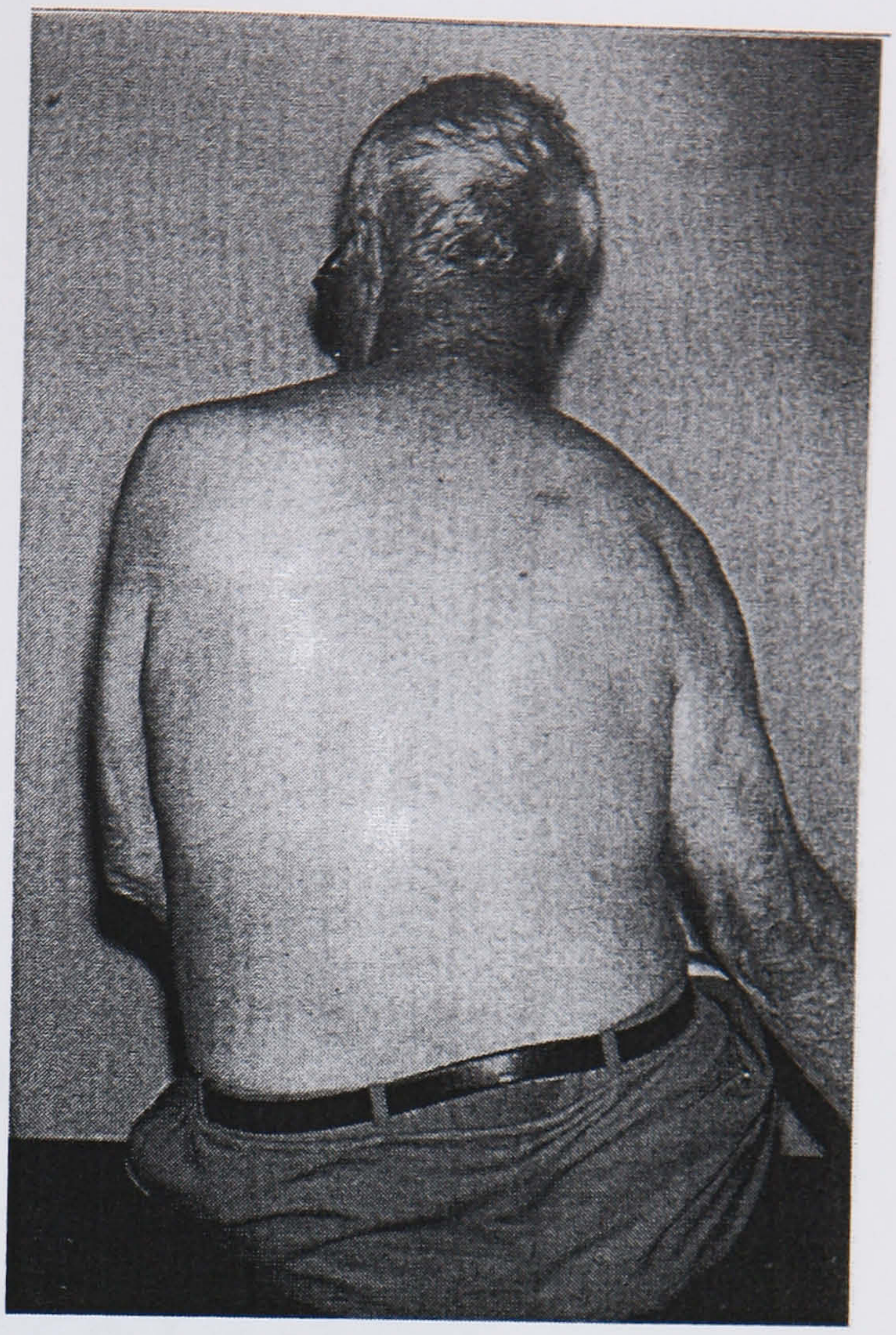
Decreased Activity / Difficulty Recruiting
No active movement just leans
Bottom recruiting – No extensors, abductors, low tone
Can't recruit extensor activity on Weak side to stabilise position

6. Other issues not included in other categories

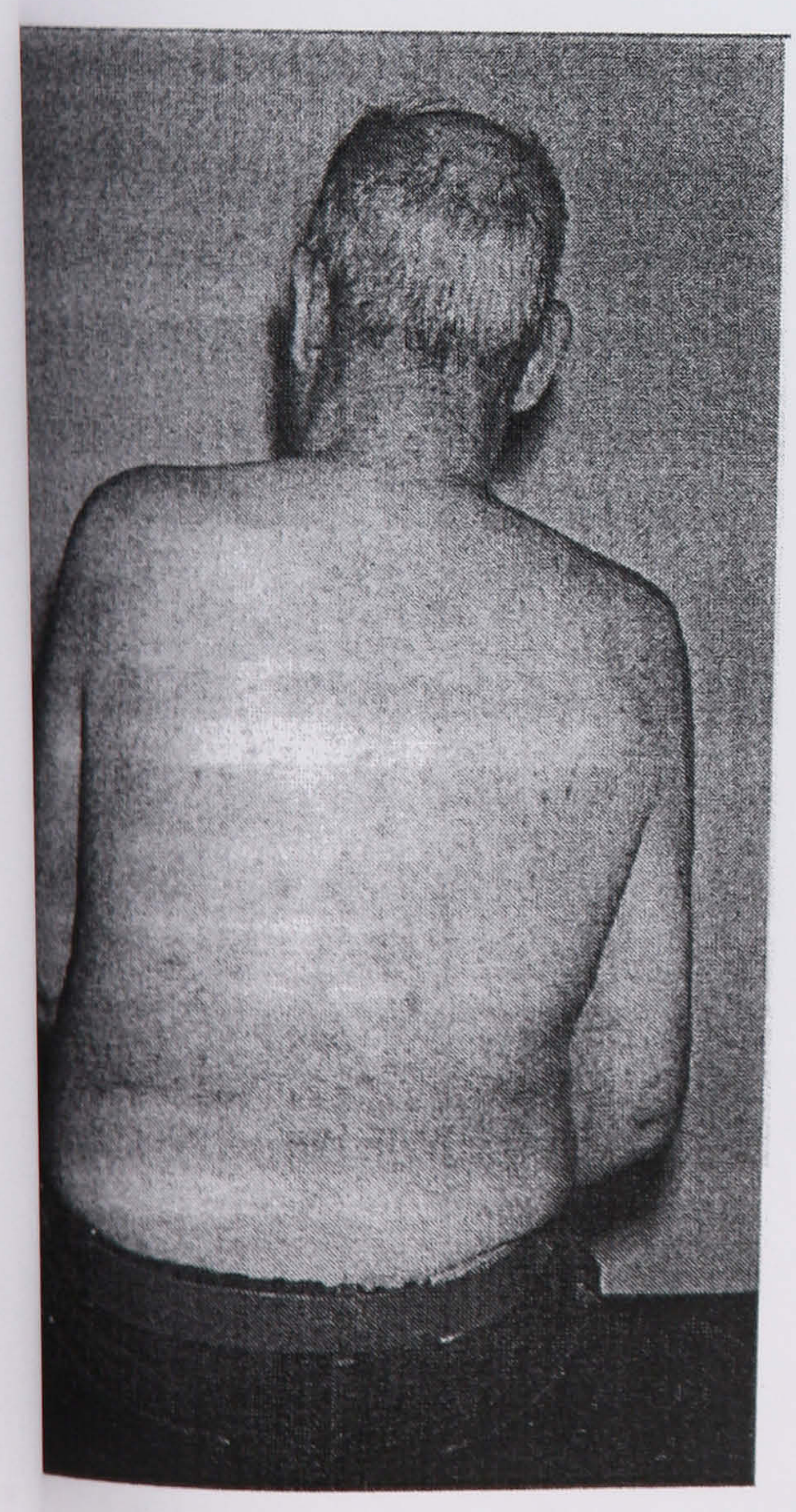
Rib cage position
Changes with time, sustainability
Face -? fixing eyes to compensate
Facial droop
Proprioception, sensory awareness, neglect
Proprioception / sensation
Sensation
Proprioception through ischial tuberosities, ?aware of weight bearing
Sensation – light touch, proprioception
Vision, neglect, perception, centre of gravity, hemianopia
Breathing control and speech – effect of posture and tone
Skin condition
Premorbid changes
Exercise tolerance
Alertness / participation
Facial expression – communication, concentration, participation, effort
How holds himself
Moved as a block
? Posture long standing
Influence of gravity on: Spatial awareness
Influence of gravity on: amount of support needed
? Changes over time
Feet - Colour swelling
Feet – swelling, shortening,

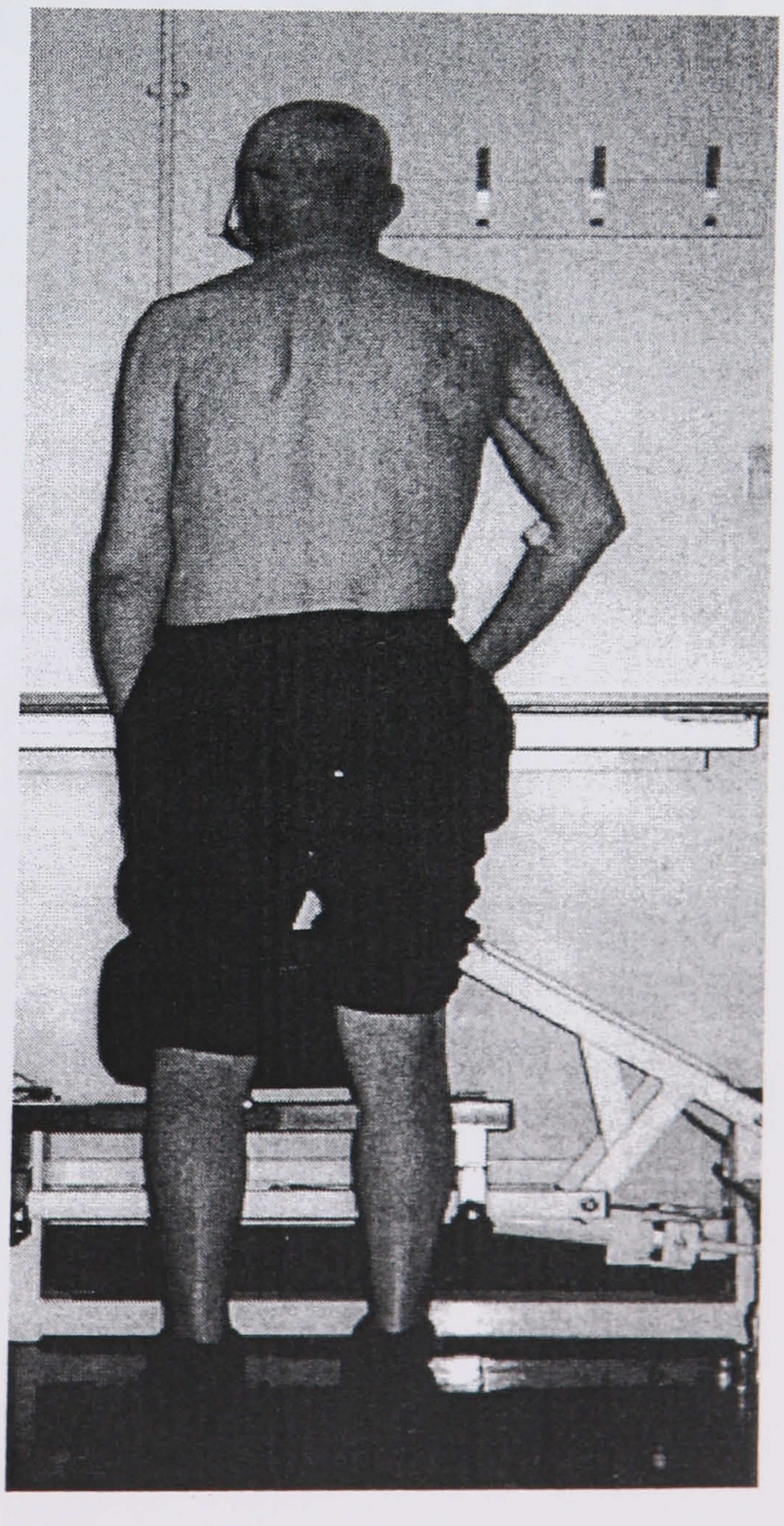
Appendix II
Photographs used in the focus groups





a) normal sitting posture (above)
a) weight transferred to the weak side (top right)
b) weight transfer to the sound side (below)
c) standing (bottom right)





Appendix III

The Brunel Balance Assessment Testing Manual

II.a Introduction

The Brunel Balance Assessment (FBA) is designed to assess functional balance for people with a wide range of abilities following stroke. There are three sections to the assessment- sitting, standing and stepping, which can be used individually or together. Each section is divided into several levels of increasingly demanding balance ability, ranging from assisted balance to moving within the base of support. At each level there is a functional performance test which gives a more sensitive test of this is level of balance.

To test a patient:

The patient performs each test in turn until they get to the level that is the limit of their abilities. For each test there is a minimal level of performance to `pass' at that level. If the patient has been unable to achieve this minimal level after three attempts then testing should cease. The levels and the functional performance tests are outlined below

II.b Testing Instructions

Level 1. Supported sitting - The subject can sit with upper limb support (i.e. taking weight through their arms) for at least 30s.

1. The subject is seated on a firm, level surface without back support and their feet flat on the floor. They can use upper limb support if they wish. Stand-by to give support if necessary.
2. Explain the test to the subject
3. Use the stop-watch to time how long they can maintain sitting balance for up to 30s. Call out the time every 10s. Instructions to Subject . "I want to time how long you can sit without me helping you. You can use your arms to support yourself if you wish. When I say GO try to keep your balance for as long as you can or until I say stop."
4. Note the time and decide whether to pass or fail. Pass if the subject keeps their balance for more than 30s. Fail if s/he keeps their balance for less than 30s, and/or requires support or supervision from the tester
5. If subject fails, make up to two more attempts

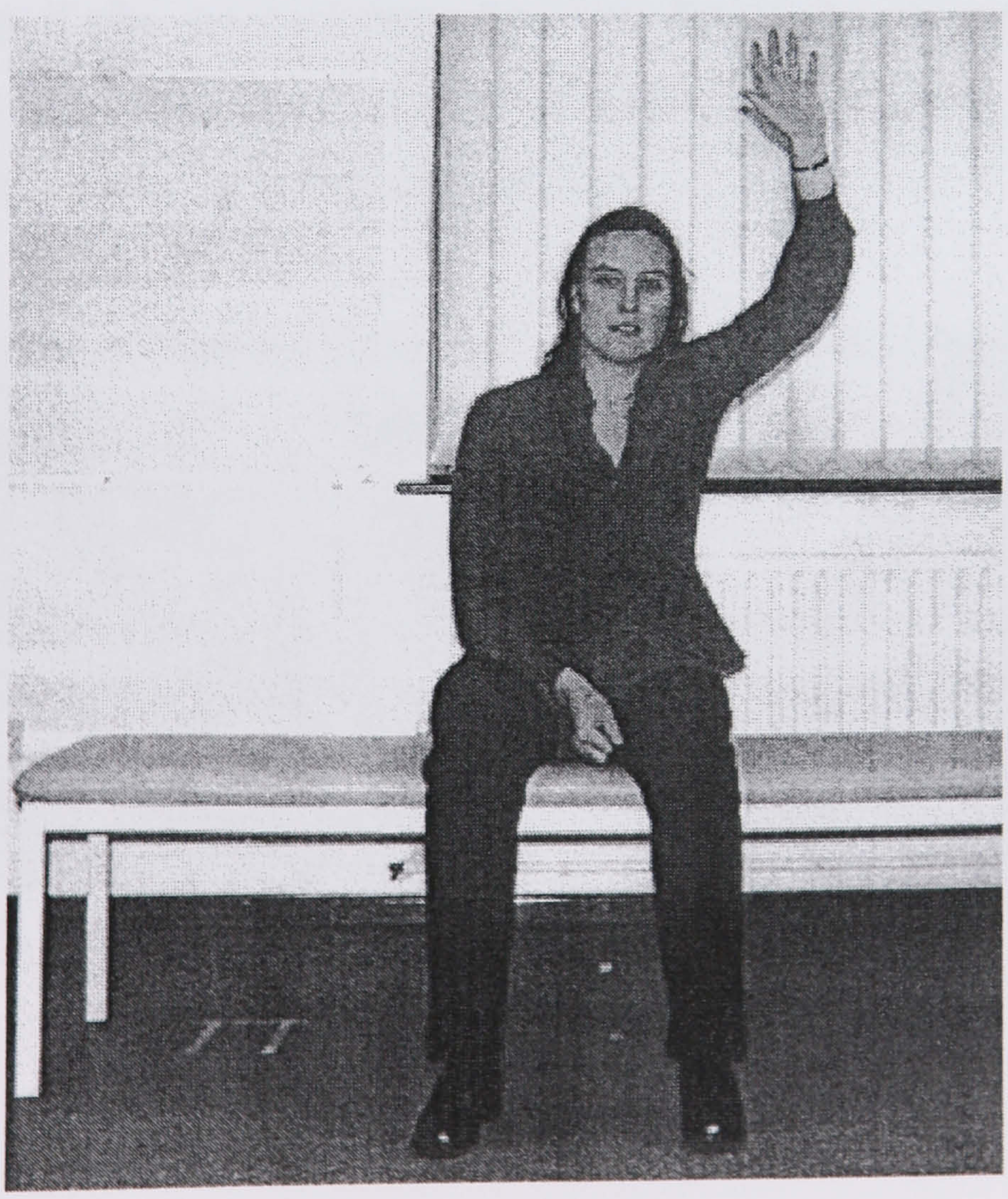
Level 2. Independent sitting balance

The subject can maintain a sitting position without upper limb support for at least 30s.

1. The subject is seated on a firm, level surface without back support or upper limb support, and their feet flat on the floor and hands resting on their lap. Stand-by to give support if necessary
2. Explain the test to the subject.
3. Use the stop-watch to time how long they can maintain sitting balance for up to 30s. Call out the time every 10s. Instructions to Subject. "I want to time how long you can sit without me helping you. Keep your arms on your lap but do not lean on them. When I say GO try to keep your balance for as long as you can or until I say stop"
4. Note the time and decide whether to pass or fail. Pass if the subject keeps their balance for more than 30s Fails if s/he keeps their balance for less than 30s, requires upper limb support, and/or requires support or supervision from the tester.
5. If the subject fails, make up to two more attempts.

Level 3. Static sitting balance - The subject can maintain the position of the base of support while moving another body segment using the *Arm Raise Test*. This is the number of times the subject can raise and lower the sound arm in 15s.

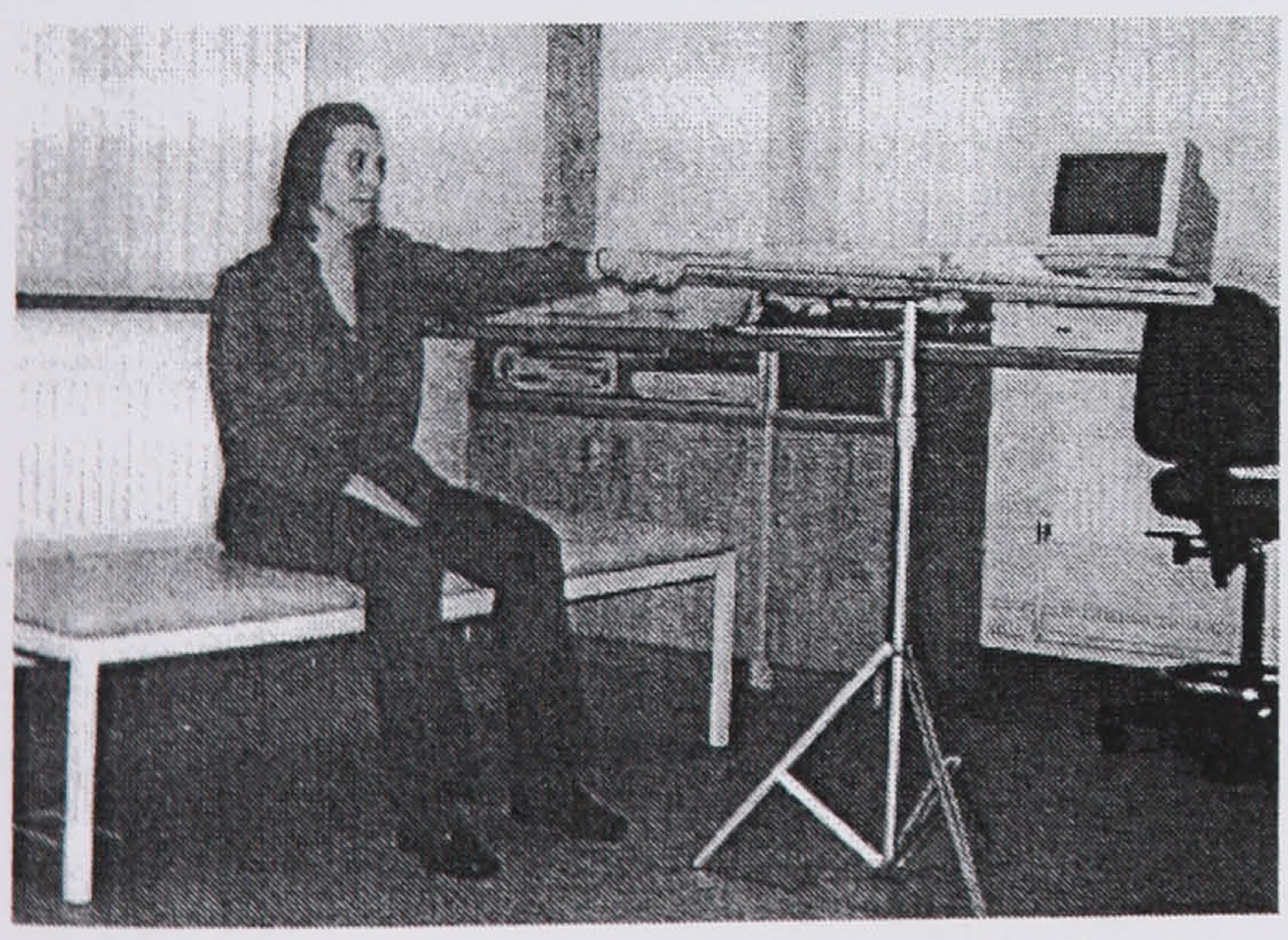
1. The subject is seated on a firm, level surface without back support, feet flat on the floor and hands resting on lap. Stand-by to give support if necessary.
2. Use the stop-watch to time 15s. Count the number of times the subject can raise their sound arm (maximum shoulder flexion) and return it back to their sound knee.
3. Explain and demonstrate the movement to the subject, get them to practise it and correct as necessary. Instructions to Subject. "I want to see how many times you can lift your sound arm in 15s. When I say GO raise and lower your arm as often as you can, until I say stop".
4. Note the score and decide whether to pass or fail. Pass if they perform more than 2 lifts. Fail if they perform less than 2 lifts. A lift does not count if the subject does not achieve full flexion (for him/her), needs to `touch down' i.e. places hand somewhere other than the sound knee to keep their balance, and/or requires support or supervision from the tester
5. If the subject fails, make up to two more attempts.

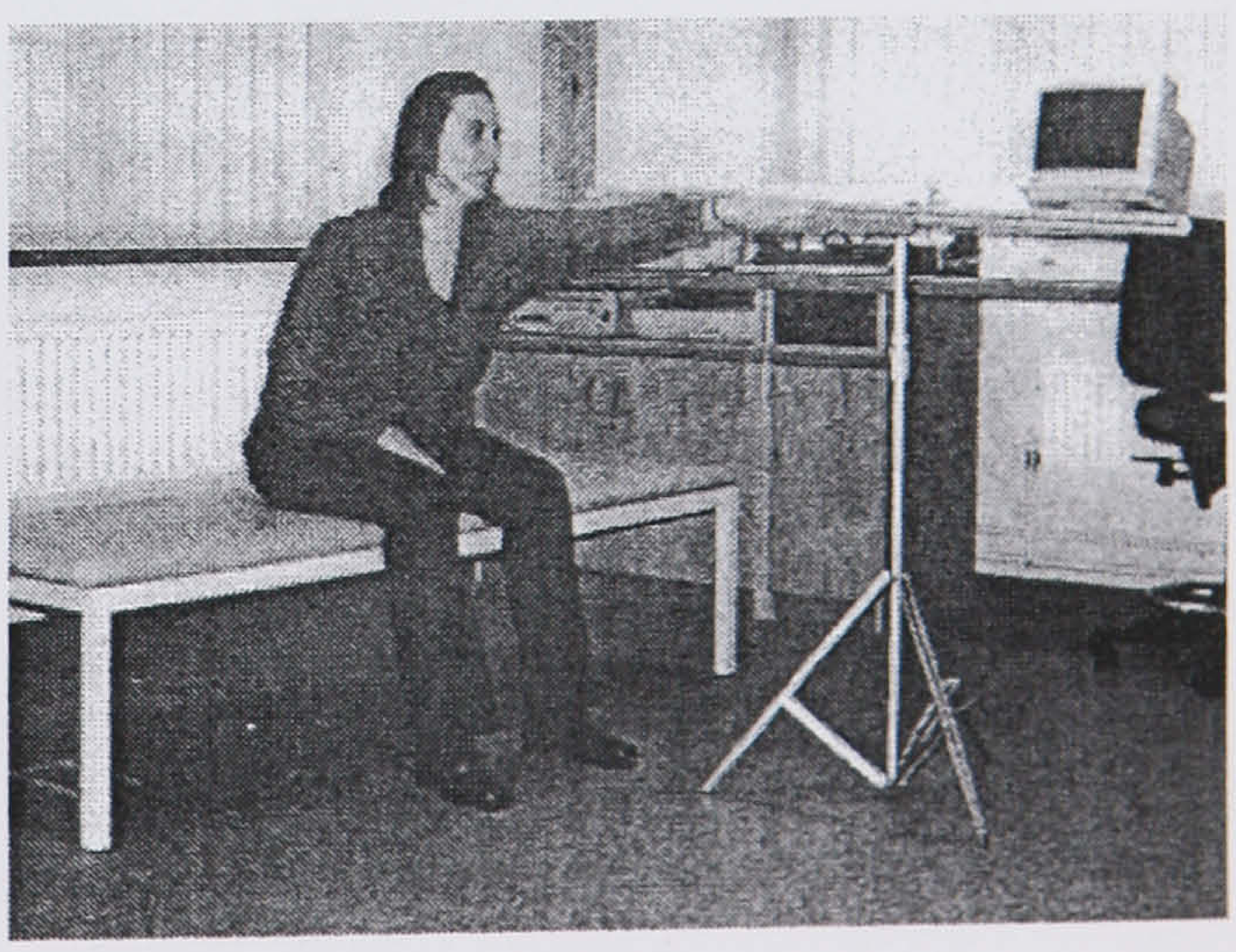


Level 4. Dynamic sitting balance

- The subject can move to limits of stability within the base of support using the *Forward Reach Test*. The distance the subject can reach forward beyond arm's length is measured.

1. The subject is seated with hips at 90 degrees on a firm, level surface without back support, feet flat on the floor and hands resting on lap. Stand-by to give support if necessary. The height of the ruler is adjusted so that it is at the level of the acromion of the sound shoulder. The subject lifts his/her sound arm to shoulder height with fingers curled into a fist while sitting in a normal, comfortable position. Position the ruler so that the end of the ruler touches the knuckles and it continues in a forward direction. The subject reaches forwards as far as possible with their hand level with the ruler. When at maximum reach the tester reads the position of the knuckle of the middle finger from the ruler.
2. Explain and demonstrate the movement to the subject, get them to practice it and correct as necessary. Instructions to Subject. "I want you to reach forwards as far as you can, keeping your hand level with the ruler. When you are at full stretch hold the position for a few seconds while I read the ruler then sit back. Keep your feet on the ground, and your bottom on the seat, do not use your weak arm for support."
3. Read the position of the knuckle of the middle finger on the ruler. Decide whether to pass or fail. Pass if s/he can reach more than 5cm. Fail if s/he can not reach 5cm, and/or requires upper limb support and/or assistance from the tester.
4. If the subject fails, make up to two more attempts





Standing Section

Level 5. Supported standing - The subject can stand with upper limb support, holding on to furniture for up to 30s.

1. The subject stands on a firm, level surface in normal shoes with feet in a comfortable, level position, and holding on to furniture if necessary. Provide support at waist height in front or to the sound side - e.g. a plinth, walking frame or back of a chair. Stand-by to give assistance as necessary.
2. Explain the test to the subject.
3. Use the stop-watch to time how long they can maintain standing balance for up to 30s. Call out the time every 10s. Instructions to Subject. "I want to time how long you can stand without me helping you. You can hold on if you wish. When I say GO try to keep your balance for as long as you can or until I say stop."
4. Note the time and decide whether to pass or fail. Pass if s/he keeps their balance for more than 30s. Fail if they keep their balance for less than 30s, and/or requires support or supervision from the tester
5. If the subject fails, make up to two more attempts

Level 6. Independent standing balance - The subject can maintain a standing position without upper limb support for up to 30s.

1. The subject stands on a firm, level surface with feet level and approximately hip distance apart without upper limb support. Stand-by to give support
2. Explain the test to the subject
3. Use the stop-watch to time how long they can maintain sitting balance for up to 30s. Call out the time every 10s. Instructions to Subject. "I want to time how long you can stand without me helping you. Keep your arms by your sides. When I say GO try to keep your balance for as long as you can or until I say stop"
4. Note the time and decide whether to pass or fail. Pass if s/he keeps their balance for more than 30s. Fails if they keep their balance for less than 30s, requires upper limb support, +/or assistance or supervision from the tester, or steps to maintain balance
5. If the subject fails, make up to two more attempts

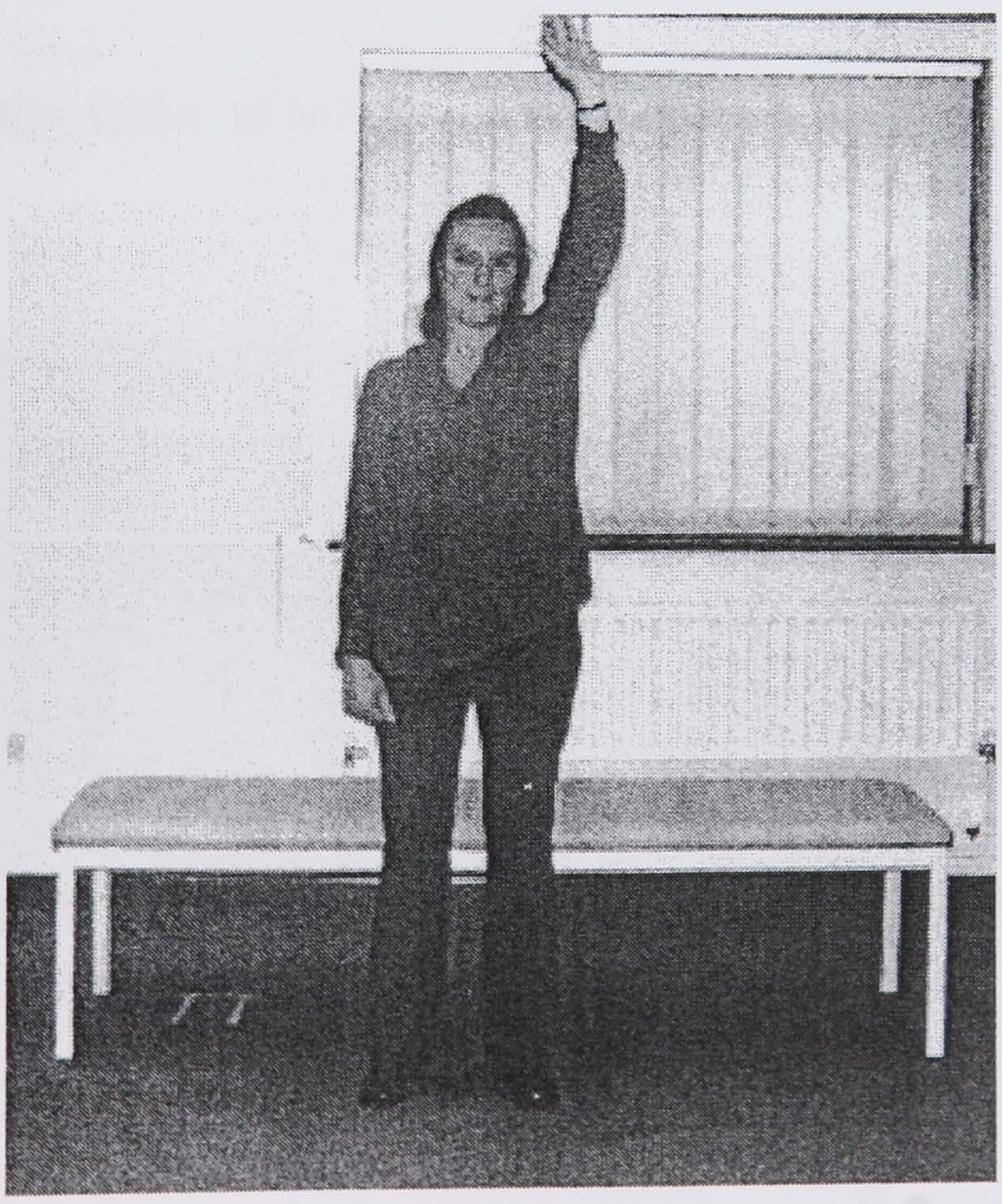
Level 7. Static standing balance - The subject can maintain their position within base of support while moving another body segment using the *Arm Raise Test* - the number of times the subject can raise and lower the sound arm in 15s is counted.

1. The subject stands on a firm, level surface with feet level without upper limb support. Stand-by to give support. Explain and demonstrate the movement to the subject, practice and correct as necessary.

3. Use the stop-watch to time 15s. Count the number of times the subject can raise (maximum shoulder flexion) and lower their sound arm (to their side) in this time. Instructions to subject. “I want to count how many times you can lift your sound arm in 30s. When I say GO raise and lower your sound arm as often as you can, until I say stop".

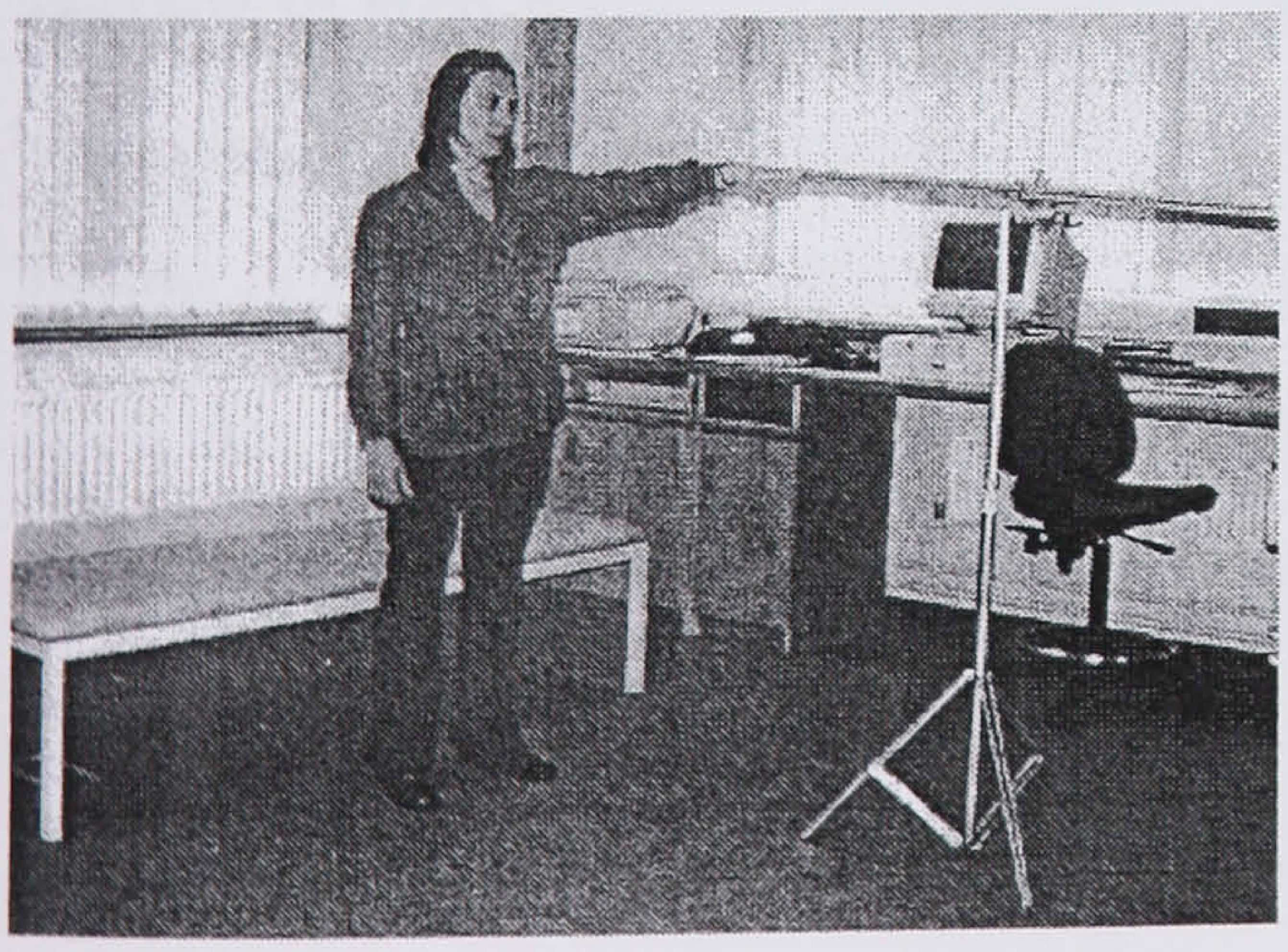
3. Note the score and decide whether to pass or fail. Pass if they perform more than 2 lifts. Fail if they perform less than 2 lifts. A lift does not count if the subject does not achieve full flexion (for him/her), needs to `touch down' i.e. places hand somewhere other than the sound knee to keep their balance, and/or requires support or supervision from the tester.

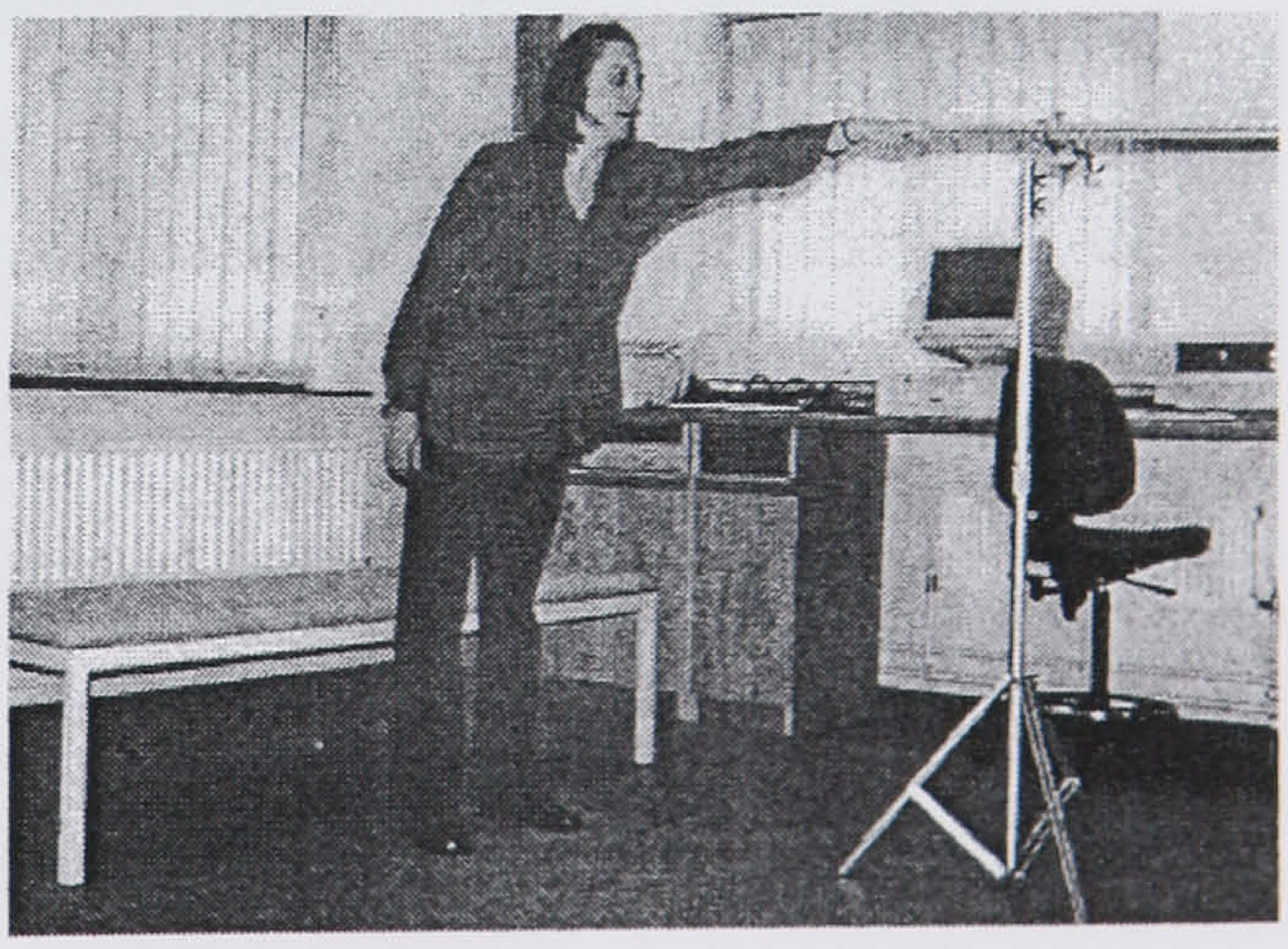
4. If the subject fails, make up to two more attempts.



Level 8. Dynamic standing balance - The subject can move to their limits of stability within the base of support using the *Forward Reach Test*. The distance the subject can reach forward beyond arm's length is measured. Pass = reach of 2cm.

1. The subject stands on a firm, level surface with feet level without upper limb support. Stand-by to give support.
2. The height of the ruler is adjusted so that it is at the level of the acromion of the sound shoulder. The subject lifts his/her arm to shoulder height with fingers curled into a fist. Position the ruler so that the knuckles are level with the end of the ruler, and the ruler points forwards in front of the subject. The subject reaches forwards as far as possible with their hand level with the ruler. When at maximum reach the tester reads the position of the knuckle of the middle finger from the ruler.
3. Explain and demonstrate the movement to the subject, get them to practice and correct as necessary. Instructions to subject. "I want you to reach forwards as far as you can with your hand level with the ruler. When at full stretch hold the position for a few seconds while I read the ruler, then return to upright. Keep your heels on the ground and do not use your weak arm for support."
4. Note the ruler reading and decide whether to pass or fail. Pass if s/he can reach more than 2cm. Fail if they can not reach 2cm, or requires upper limb support or assistance from the tester.
5. If the subject fails, make up to two more attempts.





Stepping Section

Level 9. Static step standing - The subject maintain a step-standing position without upper limb support for up to 30s.

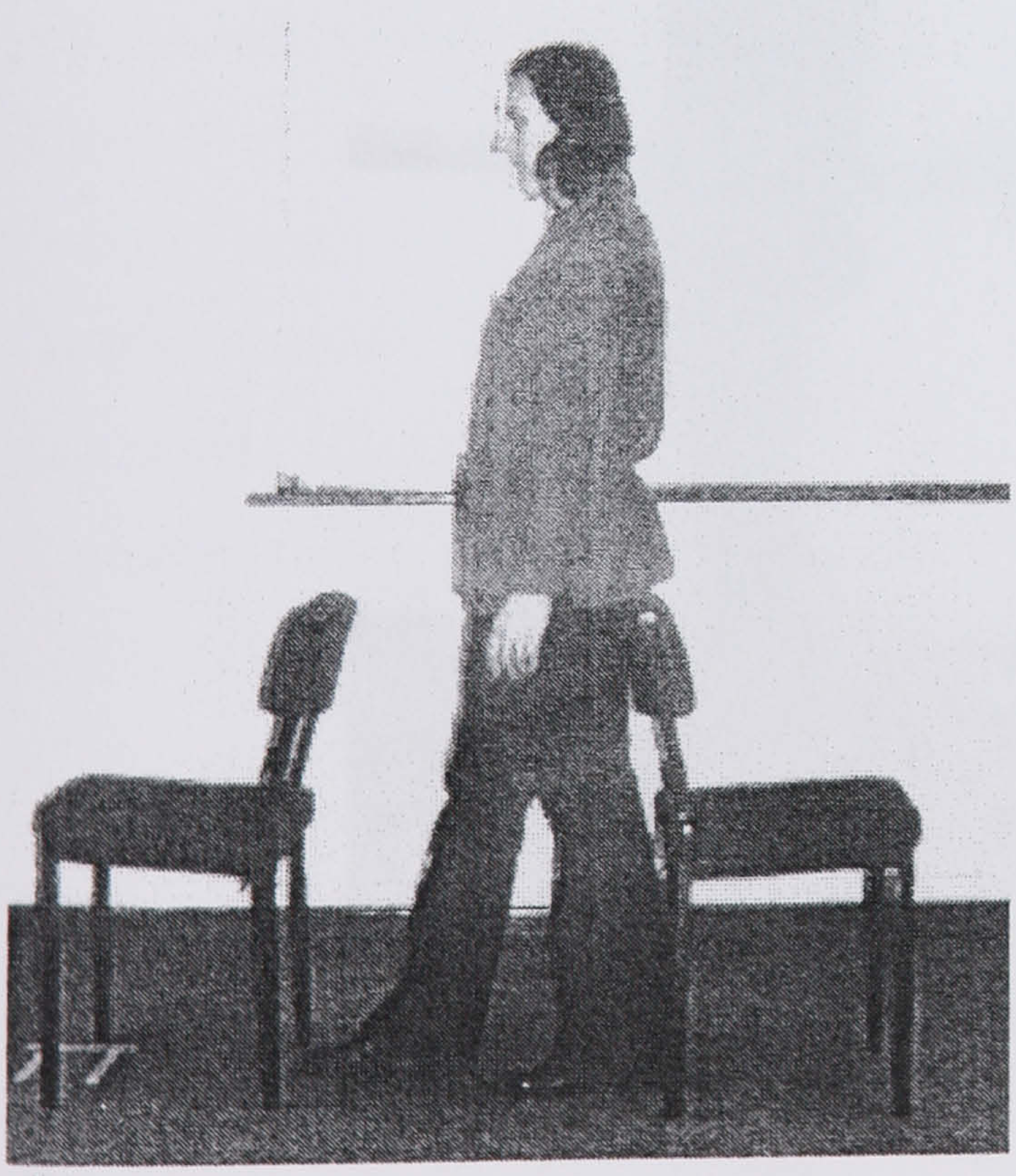
1. The subject stands without upper limb support on a firm, level surface in step standing position (sound foot in front of weak foot, with sound heel level or beyond the weak toes, both knees extended).
2. Explain the test to the subject, demonstrate and practise as necessary. Stand-by to give support.
3. Use the stop-watch to time how long they can maintain the step standing position for up to 30s. Call out the time every 10s. Instructions to Subject . "I want to time how long you can stand without me helping you. Keep your arms by your sides. When I say GO try to keep your balance for as long as you can or until I say stop"
4. Note the time and decide whether to pass or fail. Pass if s/he keeps their balance for more than 30s. Fail if s/he keeps their balance for less than 30s, requires upper limb support, and/or support or supervision from the tester
5. If the subject fails, make up to two more attempts.

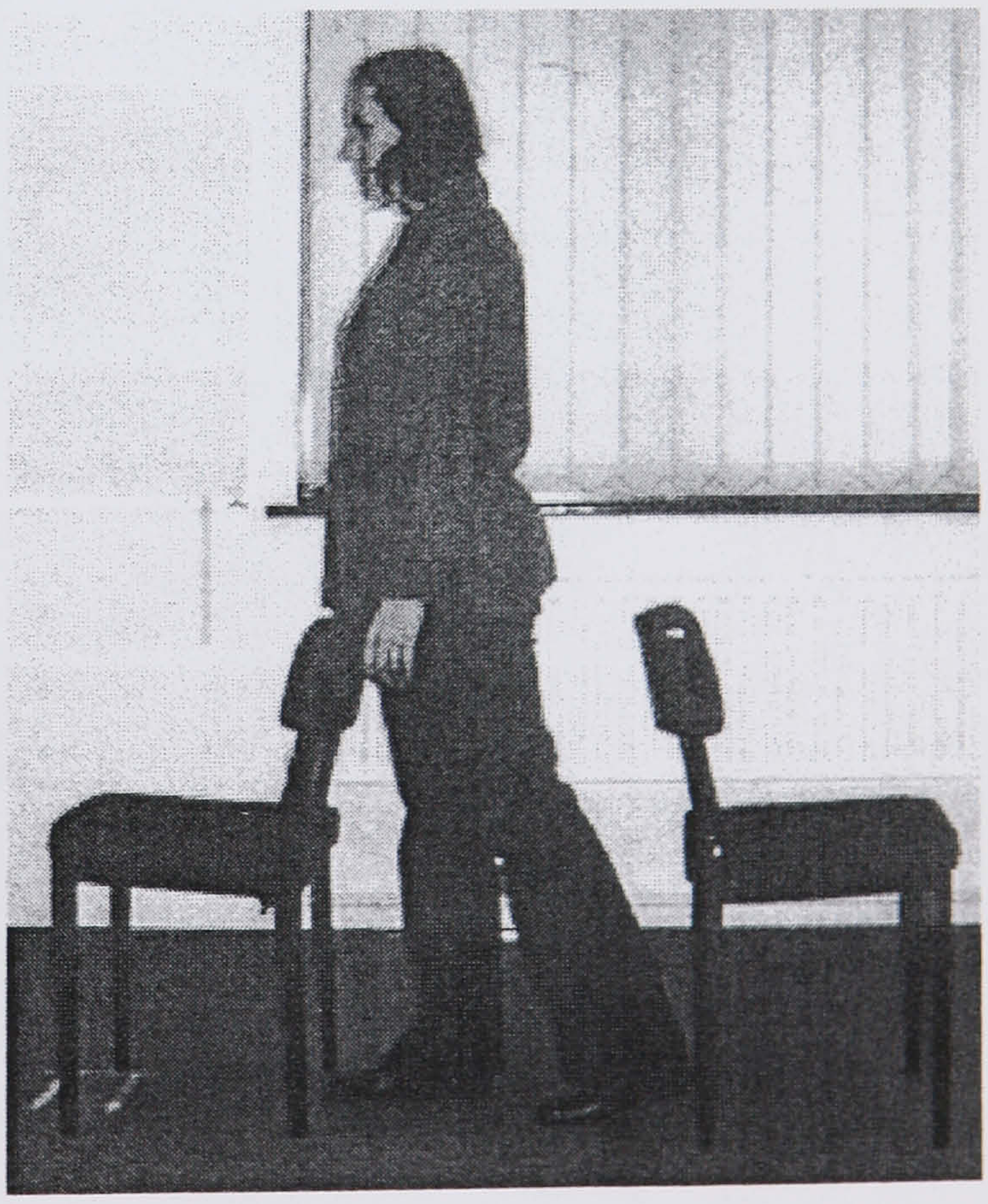
Level 10. Stepping within BoS - The subject can move within the base of support in a step-standing position using the *Weight Shift Test* - counting the number of times s/he can transfer their weight onto the weak leg in 15s.

1. Starting position The subject stands without upper limb support on a firm, level surface in step-standing position (weak foot in front, with weak heel level or beyond the sound toes). A perching stool or walking frame (or similar) adjusted to hip/tummy height is positioned so that the horizontal bar is over the 5th metatarsal of the weak foot. Another frame or stool is positioned behind the subject at hip/ bottom level, so their bottom touches the stool when their weight is on the sound leg. Stand-by to give support as necessary.

   Movement The subject transfers their weight on to the weak leg so that their tummy touches the back of the stool, and then back on to the sound leg so that their bottom touches the other stool. They need to stand upright and keep their hips neutral/extended. The sound heel may lift as weight is transferred forwards but it must be on the floor when weight bearing.

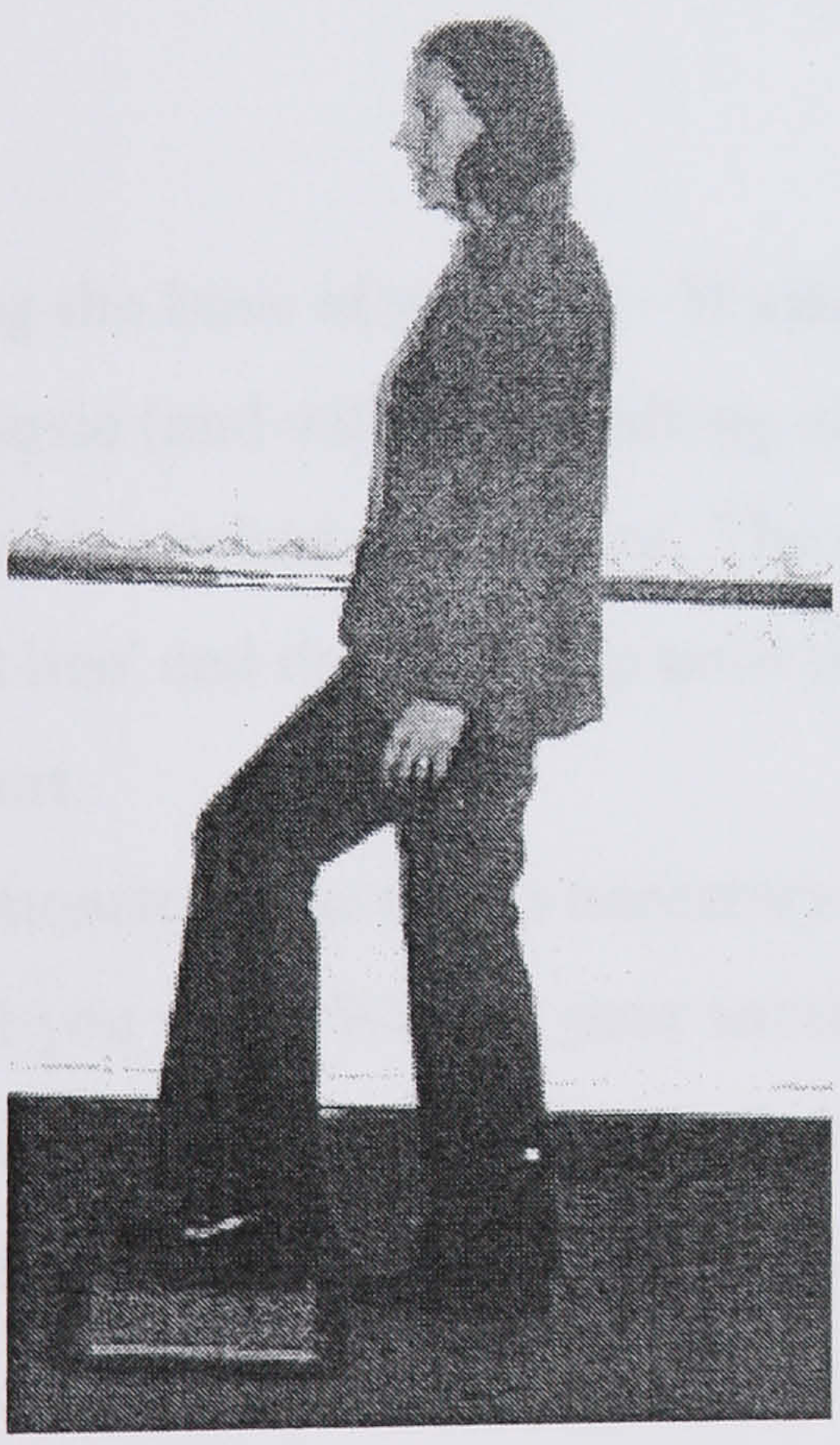
2. Explain and demonstrate the test, practise and correct as necessary. <u>Instructions to Subject</u> " I want to count how many times you can transfer your weight from one leg to the other, and back again. When I say GO transfer your weight on to the weak leg so that your tummy touches the stool/frame, then back on to the sound leg so that your bottom touches the other stool/frame. Keep your hips and knees straight when your weight is on the leg, but your can bend your sound knee and raise your heel as you bring your weight forwards. Do this as many times as you can until I say stop."
3. Use the stop-watch to time 15s, and count the number of times they touch the bar of the frame at the front, i.e. how often they transfer their weight on to the weak foot.
4. Note the score and decide whether to pass or fail. <u>Pass</u> if they performs more than 2 transfers. <u>Fail</u> if they perform less than 2 transfers. A transfer does not count if the subject does not touch the stool/frame at front or back, uses upper limb support, and/or requires support or supervision from the tester
5. If subject fails, make up to two more attempts





Level 11. Dynamic single stance - The subject can maintain single stance on the weak leg while moving the other leg using the *<u>Step-Tap Test</u>* - counting the number of times the subject can place his sound leg on and off a step while standing on the weak leg in 15s.

1. Subject stands on a firm level surface with feet level. A 7.5-10cm block is positioned a hands width (10cm) in front of his/her toes. The subject places his/her sound foot on and off the block as often as possible within 15s (but does not step up). S/he should place his/her whole foot on the block. Stand-by to give support.
2. Explain and demonstrate the test, practise and correct as necessary.
3. Use the stop-watch to time 15s, and count aloud the number of steps they perform.
4. Note the score and decide whether to pass or fail. <u>Pass</u> if they performs more than 2 steps. <u>Fail</u> if they performs less than 2 steps. A step does not count if the subject uses upper limb support, and/or requires support or supervision from the tester.
5. If the subject fails, make up to two more attempts.



Level 12. Changing the base of support with support - Walking with an aid. Subject can walk without assistance (but may use a walking aid) using the *5m Walk Test with an aid*.

1. A distance of 5m is marked on the floor. The subject starts to walk a couple of strides before the `start line' and does not stop until they have crossed the `finish line'. Stand-by to give support.
2. Explain and demonstrate the test as necessary. Instructions to the subject. "I am going to time how fast you walk. Walk at your natural pace between these two markers. Do not slow down until you have crossed the finish line. Start when I say GO."
3. Use the stop-watch to time how long it takes to walk this distance.
4. Note the time and decide whether to pass or fail. Pass if s/he completes the distance in less than 1 minute. Fail if they take more than 1 minute or if physical or verbal support is needed.
5. If the subject fails, make up to two more attempts

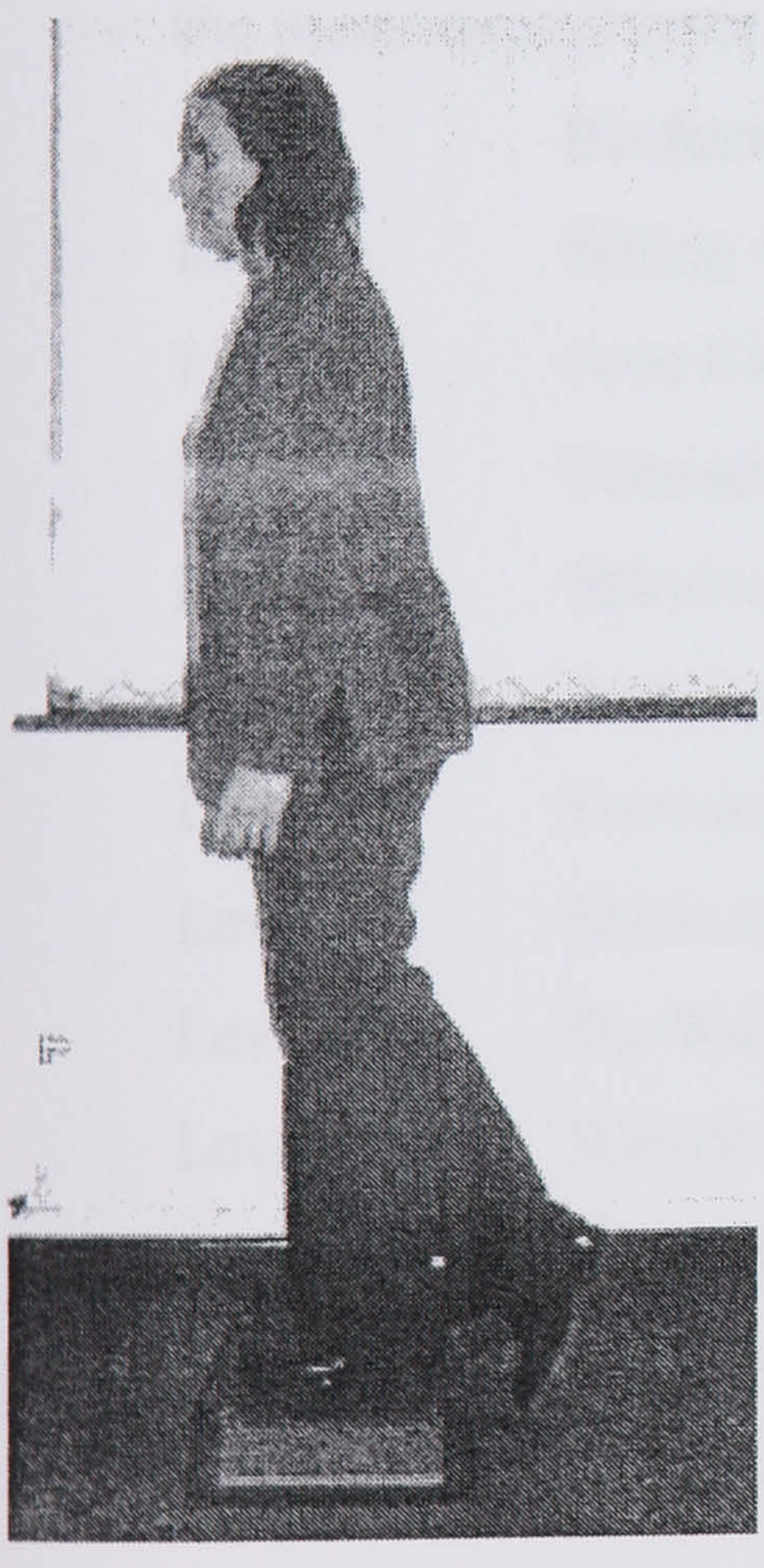
N.B. If the subject is already able to walk without an aid, pass this level and progress to the next.

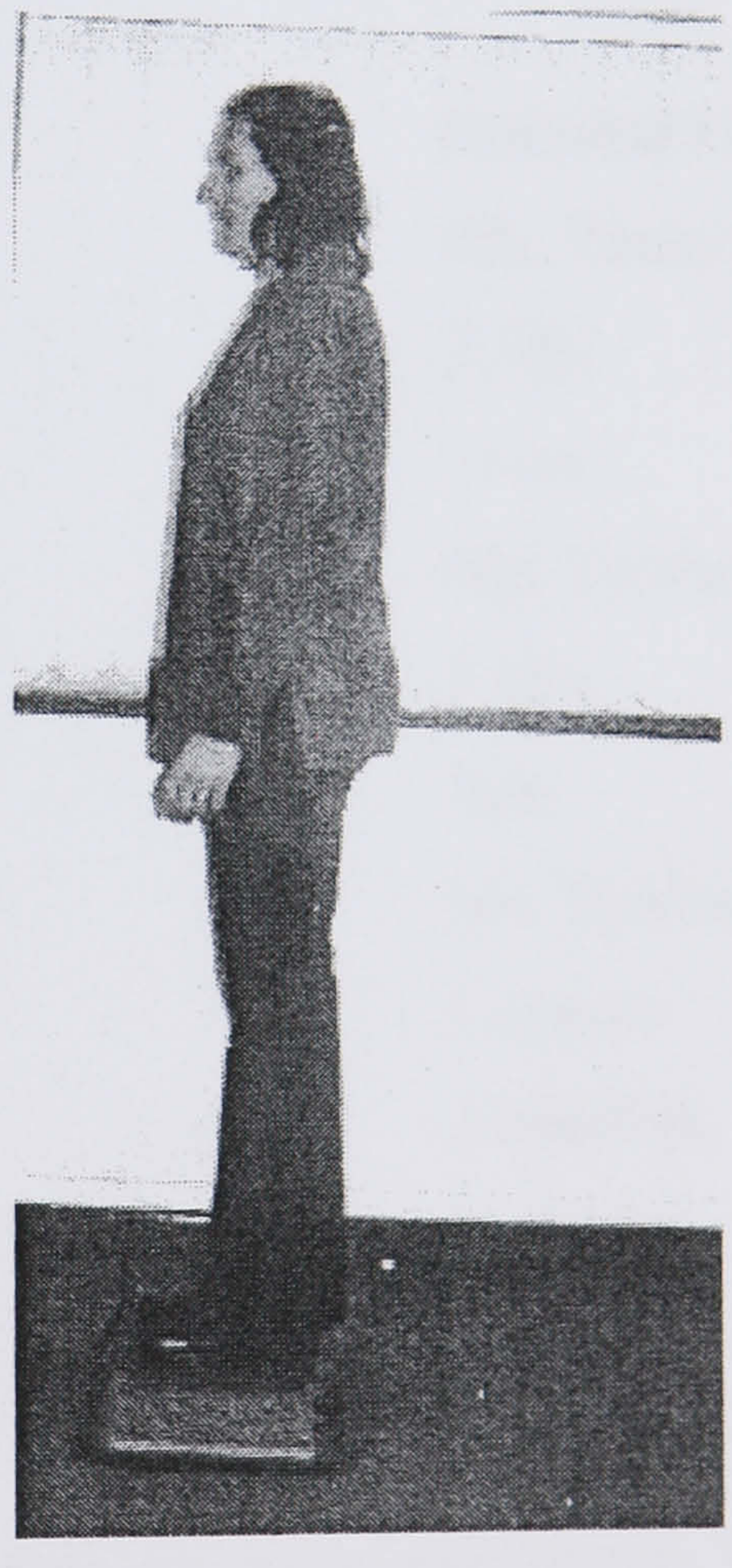
Level 13. Changing the base of support - Walking without an aid. The subject can walk without assistance (and without a walking aid) using the *5m Walk Test*.

1. A distance of 5m is marked on the floor. The subject starts to walk a couple of strides before the `start line' and does not stop until they have crossed the `finish line'. Stand-by to give support.
2. Explain and demonstrate the test as necessary. Instructions to the subject. "I am going to time how fast you walk. Walk at your natural pace between these two markers. Do not slow down until you have crossed the finish line. Start when I say GO."
3. Use the stop-watch to time how long it takes to walk this distance.
4. Note the time and decide whether to pass or fail. Pass if s/he completes the distance in less than 1 minute. Fail if it takes more than 1 minute, or if physical or verbal support is needed.
5. If the subject fails, make up to two more attempts

Level 14. Changing the Base of Support - Stepping. The subject can change their base of support using the *Step-Up Test,* leading with their weak leg.

1. Subject stands on a firm level surface with feet level. A 7.5/10cm wooden block is positioned a hands width (10cm) in front of his/her toes. The subject steps up, onto, and off the block leading with their weak leg as often as possible within 15s. A Step-up is completed when weak leg is placed on the floor again. Stand-by to give support.
2. Explain and demonstrate the test, practise and correct as necessary. Instructions to Subject "I want you to step up on to the block and then off again, leading with your weak leg. When I say GO do this as often as you can until I say stop."
3. Use the stop-watch to time 15s, and count the number of steps-up performed.
4. Note the score and decide whether to pass or fail. Pass if they can perform 1 step-up. Fail if they cannot complete 1 step-up and down, or support is required.
5. If subject fails, make up to two more attempts.





II.c Changes made as a result of the scale development.

1. **Hierarchy of the ordinal scale**

The order of the stepping section was altered to:

Level 9. Static step standing for 30s *(Yes/no)*

Level 10. Supported change of the base of support *(5m Walk Test with an aid)*

Level 11. Stepping within base of support *(Weight Shift Test)*

Level 12. Changing the base of support - *(5m Walk Test without an aid)*

Level 13. Dynamic single stance *(Step Test with weak leg)*

Level 14. Changing the Base of Support – stepping *(Step-Up Test)*

2. **Redundancy**

Levels 2 and 6 (static balance without upper limb support in sitting and standing respectively) were removed.

3. **Measurement error**

The minimal pass score for each item/level of the scale was adjusted according to the measurement error.

Level	Performance Test	Minimal Pass
Level 1.	Sitting with upper limb support	30s, Yes/no
Level 2.	Arm Raise Test in sitting	3 lifts
Level 3.	Forward Reach Test in sitting	11cm
Level 4.	Standing with upper limb support	30s, Yes/no
Level 5.	Arm Raise Test in standing	3 lifts
Level 6.	Forward Reach Test in standing	7cm
Level 7.	Static step standing	30s, Yes/no
Level 8.	5m Walk Test with an aid	1 minute
Level 9.	Weight Shift Test	3 transfers
Level 10.	5m Walk Test without an aid	1 minute
Level 11.	Step Test with weak leg	2 taps
Level 12.	Step-Up Test	1 step-up

4. **To reduce random error** the ratio measures (forward reach tests, and the 5m walk test) should be measured twice, and mean values calculated. These tests and the Weight Shift test should be clearly demonstrated and practised before measurements are taken.

II.d The final score sheet

Subject name:

Level	Score	Pass/Fail
1. Supported Sitting (with UL support 30s)		Y/N
2. Static sitting (Arm Raise Test)		Pass=3 lifts
3. Dynamic sitting (Forward Reach Test)		Pass=11cm
4. Supported standing (30s)		Y/N
5. Static standing (Arm Raise Test)		Y/N
6. Dynamic standing (Forward Reach Test)		Pass=7cm
7. Static step standing (30s)		Y/N
8. Supported single stance (5m Walk Test with an aid)		Pass=1min
9. Dynamic double stance (Weight Shift Test)		Pass=3 shifts
10. Changing the BoS (double/ single stance 5m Walk Test)		Pass=1min
11. Dynamic single stance (TapTest)		Pass=2 steps
12. Changing the BoS (Step-Up Test)		Pass=1step-ups.

Appendix IV
Testing manual for the 'other tests'

III.a Other Tests of Balance

III.a.i Berg Balance Scale

The Berg Balance Scale consists of fourteen tasks, each of which is divided into four levels, and the tester scores the lowest level, which applies for the subject's performance. The test is not hierarchically arranged and the tasks are not designed to be used individually (Berg et al 1989a, 1995; Wood-Dauphinee, et al 1996; Stevenson & Garland 1996). The subject attempts each task in turn, and his/her performance on each is categorised. One attempt at each task is allowed unless stated otherwise. The score for each task is summed to give a total score for the test.

1. Sit to Stand

Instruction; Please stand up, try not use your hands for support

4) Able to stand, no hands and stabilises independently

3) Able to stand independently using hands

2) Able to stand using hands after several tries

1) Needs minimal assistance to stand or stabilise

0) Needs moderate /maximum help to stand

2. Standing unsupported

Instruction: Stand for two minutes without holding on.

4) Able to stand safely for 2 mins

3) Able to stand safely for 2 mins with supervision

2) Able to stand for 30s unsupported

1) Needs several tries to stand for 30s unsupported

0) Unable to stand for 30s unassisted

If subject can complete this then miss the sitting section and score 4.

3. Sitting unsupported, feet on floor

Instruction: Sit with arms folded for two minutes

4. Able to sit safely and securely for 2 minutes

3. Able to sit safely for 2 minutes with supervision

0. Able to sit for 30s

1. Able to sit for 10sec

2. Unable to sit without support for 10s.

4. Standing to Sit

Instruction: Please sit down

4. Sits safely with minimal use of the hands
3. Controls descent by using hands
2. Uses back of legs against the chair to control descent
1. Sits independently but has uncontrolled descent
0. Needs assistance to sit down

5 Transfers

Instruction: Please move from bed to chair and back again. One way towards a seat with arm rests and one way to a seat without armrests.

4. Able to transfer safely with only minor use of hands
3. Able to transfer safely with definite need of hands
2. Able to transfer with verbal cueing and/or supervision
1. Needs one person to assist
0. Needs two people to assist

6. Standing unsupported with eyes closed

Instruction. Please close your eyes and stand still for 10s.

4. Able to stand for 10s safely
3. Able to stand for 10s with supervision
2. Able to stand for 3s
1. Unable to keep eyes closed for 3s but stays steady
0. Needs help to keep from falling

7. Standing unsupported with feet together

Instruction: Place your feet together and stand without holding on.

4. Able to place feet together independently and stand for 1min
3. Able to place feet together independently and stand for 1min with supervision
2. Able to place feet together independently but unable to hold for 30s
1. Needs help to attain position but able to stand for 15s with feet together
0. Needs help to attain position and unable to maintain for 15s

The following items are performed while standing unsupported.

8. Reaching forwards without stretched arm

This is the same as the Forward Reach Test - See Appendix II (Brunel Balance Assessment) for details of how to perform it.

4. Can reach confidently more than 25.5 cm (10 inches)
3. Can reach more than 13cm (5 inches)
2. Can reach forwards less than 5cm (2 inches)
1. Reaches forwards but requires supervision
0. Needs help to keep falling

9. Pick up an object from the floor

Instruction: Pick up the beanbag that is placed in front of you feet.

4. Able to pick up the beanbag easily and safely
3. Able to pick up the beanbag but requires supervision
2. Unable to pick it up but reaches to within 1-2 inches and keeps balance independently
1. Unable to pick it up and needs supervision while trying
0. Unable to try/ or needs assistance to keep from falling

10. Turning to look over shoulder

Instruction: Turn your head to look behind you. Repeat to the other side

4. Looks behind to both sides and shifts weight well
3. Looks behind to one side only other side shows less weight shift
2. Turns head sideways only but maintains balance
1. Needs supervision when turning
0. Needs assistance to keep from falling

11 Turn 360 degrees.

Instruction: Turn around completely in a full circle. Then return the other way.

4. Able to turn 360 safely in less than 4s to each side
3. Able to turn safely to one side only and in less than 4s
2. Able to turn 360 slowly but safely
1. Needs supervision or verbal cueing
0. Needs assistance while turning

12 Stool touch

Instruction: Place each foot alternately on the stool. Continue until each foot has touched the stool four times

4. Able to complete 8 steps safely and independently in 20s

3. Able to complete 8 steps in more than 20s

2. Able to complete 4 steps without aid, with supervision

1. Able to complete 2 steps with minimal assistance

0. Needs assistance to keep from falling/ Unable to perform

13. Standing Unsupported, one foot in front

Instruction (demonstrate): Place one foot directly in front of the other. If you feel that you can not place your foot directly in front then try to step ahead so that the heel of your front foot is ahead of the toes of your back foot.

4. Able to place foot in tandem position independently and hold for 30s

3. Able to place foot ahead of other independently and hold for 30s

2. Able to take small step independently and hold for 30s

1. Needs help to step but can hold for 15s

0. Loses balance when stepping or standing

14. Standing on one leg.

Instruction: Stand on one leg as long as you can without holding on.

4. Able to lift leg independently and hold for 10s

3. Able to lift leg independently and hold for 5-10s

2. Able to lift leg independently and hold for 3+ sec

1. Tries to lift leg unable to hold for 3s but remains standing independently

0. Unable to try or needs assistance to prevent fall.

Total score /56

III.a.ii Motor Assessment Scale (Sitting Section)

The Motor Assessment Scale consists of six tasks which the subject either passes or fails, giving a maximum score of six. (Carr et al 1985; Poole & Whitney 1988; Loewen & Anderson 1988; Dean & Mackey 1992; Nitz & Gage 1995)

1. Sits with UL support only.
2. Sits without UL support for 10s.
3. Sits unsupported with weight well forwards and evenly distributed. Weight should be well forwards at the hips, head and thoracic spine extended, weight evenly distributed on each side.
4. Sitting, turns head and trunk to look behind.
5. Sitting, reaches forwards to touch floor, and returns
6. Sits on stool unsupported, reaches sideways to touch floor and returns .

III.b Tests of Related Constructs

III.b.i Motor Impairment - The Motricity Index (lower limb section)

This simple measure of motor loss/ weakness is based on the MRC Oxford scale, but weighted scores are used (Demeurisse et al 1980; Wade et al 1987; Collin & Wade 1990; Collen et al 1990). The tests for the lower limb only were used.

The patient sits in a chair or over the edge of the bed.

Ankle dorsiflexion: The weak foot is relaxed in plantarflexed position. The patient is asked to dorsiflex the weak foot. Monitor the contraction of tibialis anterior.

Knee extension: The weak foot is unsupported, knee is at 90 degrees. The patient extends his knee to touch the examiner's hand, which is level with knee. Monitor quadriceps.

Hip flexion: Patient is sitting with hips at 90 degrees. The patient is asked to lift his/her knee towards their chin. Check for and prevent trick movements. Monitor contraction of ilio-psoas.

Scoring

0 = No Movement; 9 = Palpable contraction in muscle but no movement; 14 = Movement seen but not full range or against gravity; 19 = Movement - full range against gravity but not resistance; 25 = Movement - full range, against gravity and resistance, but weaker than other side; 33 = Normal power

Motoricity Score = (a+b+c+1)/3

III.b.ii Motor Disability - Rivermead Mobility Index

This simple index consists of fifteen questions; one point is given for each positive answer with a maximum score of fifteen. It is administered verbally except an observation of standing balance (Collen et al 1990a; 1990b)

1. Rolling Do you turn over in bed without help?
2. Lie→ Sit Do you get from lying in bed to sitting over the edge of the bed on your own?
3. Sitting balance. Can you sit over the edge of the bed for 10s without holding on?

4. Sit → Stand. Can you stand up within 15s and stand without holding on for 15s (using hands and aid if necessary)?
5. Standing. Observe standing balance for 10s without aid
6. Transfers. Do you move from bed to chair and back again without help?
7. Walking inside with aid if necessary. Do you walk 10m, with an aid if necessary without standby help?
8. Stairs. Do you manage a flight of stairs without help?
9. Walking outside (uneven ground). Do you walk around outside, on pavements without help?
10. Walking inside without aid. Do you walk 10m inside without a splint, aid or standby help?
11. Picking up from the floor. If you drop something on the floor could you walk 5m, pick it up and walk back again?
12. Walking outside (uneven ground). Do you walk over uneven ground (grass, gravel, dirt snow etc) without help
13. Bathing. Do you get in/out of the bath or shower unsupervised and wash yourself?
14. Up and down four steps. Do you manage to get up and down four steps with no rail but with an aid if necessary?
15. Running. Do you run 10m without limping in 4s (fast walk is acceptable)

III.b.iii Motor Handicap - The Nottingham Extended ADL Scale

This index has four sections, each form a hierarchical scale for people with stroke. Scoring is on a 0-3 scale, with 0 indicating the subject is unable to perform the activity, and three indicates that s/he can manage the activity alone with ease, giving a maximum score of eighteen The mobility section is used here. (Nouri & Lincoln 1987, Lincoln & Gladman 1992; Wade 1992; Gladman et al 1993).

Do you?	Not at all (0)	With help (1)	Alone w. diff (2)	Alone easily (3)
Walk around outside				
Climb stairs				
Get in and out of a car				
Walk on uneven ground				
Cross roads				
Use public transport				

III.b.iv Activities of Daily Living - Barthel Index

The Barthel Index is the most widely known and used test of ADL (Wade & Collin 1988; Collin et al 1988; Wade 1992). It tests what the patient does do, not what he can do. The main aim is to establish the degree of independence from help, physical or verbal however minor or for whatever reason. The use of aids to be independent is allowed. Supervision for any reason means the patient is not independent. The patients' performance should be established from the best available source - the patient, relative/carer, or health care professionals. Observation is useful but direct testing is unnecessary. Usually performance over the last 1-2 days is considered. The middle categories imply that the patient supplies at least 50% of the effort. There is maximum score of 20.

Bowels (in the proceeding week).

0 = incontinent or needs enema

1 = occasional accident (1 x wk)

2 = continent

Bladder (in the proceeding week).

0 = incontinent, or catheterised or unable to manage alone

1 = occasional accident (within last 24hrs)

2 = continent, including complete self-management of catheterisation

Grooming (within last 24hrs). This refers to personal hygiene: cleaning teeth, fitting false teeth, combing hair, shaving, washing face. Implements can be supplied by a helper.

0 = Needs help with personal care

1 = Independent

Toilet Use includes reaching the toilet/commode, undressing, cleaning self, dressing and leaving.

0 = dependent

1 = Needs some help but can do some alone

2 = Independent

Feeding involves eating any normal food (not restricted to soft food). Food can be cooked and served by others but not cut up.

0 = Unable

1 = Needs help cutting, spreading, etc, but can feed him/herself

2.= Independent

Transfer (bed ←→ chair)

0 = Unable, no sitting balance, two people to help

1 = Major help - can sit but needs physical assistance of 1strong/skilled helper. 2 `normal' people

2 = Minor help - can be assisted easily by one person (verbally or physically)

3 = Independent, may use an aid.

Mobility refers to indoor mobility around the house or ward. An aid may be used. If using a wheelchair, corners and doors must be negoiated unaided.

0 = Immobile

1 = Wheelchair independnent, including steering and corners/ doors

2 = walks with help of 1 (verbal or physical)

3 = Independent although may use an aid

Dressing

0 = Dependent

1 = Help with buttons, zips etc (check!) but can put on some clothes unaided

2 = Independent - including buttons, zips, laces etc. Can select and put all clothes on/off although they may be adapted.

Stairs. To be independent the patient must carry any walking aids.

0 = Unable

1 = Needs help (verbal, physical, with aid)

2 = Independent

Bath/Shower.

0 = Dependent

1 = Independent - including getting in and out and washing self. If using a shower, the patient must be unsupervised and unaided.

Total ..../20

III.c. Tests of Impairments

Motor Impairment -Motoricity Index (described in section b.i)

III.c.i Sensation - Rivermead Assessment of Somatosensory Performance

The testing protocol is based on that described in the Rivermead Assessment of Somatosensory Performance - RASP (Winward et al 2000). This is a comprehensive somatosensory testing schedule but allows the tester to choose the tests which are relevant to the patient/subject. It is based on the standardisation of clinical methods of assessment and is reliable (Winward et al 2000). The following aspects of sensation were tested: surface pressure touch, surface localisation, sensory extinction, and proprioception - movement and direction discrimination and the scores for these aspects totalled to give an overall score for sensation. The upper and lower aspect of the metatarsophalangeal (MTP) area of the first (big) toe of the weak foot was tested. Subject exclusions - people with sensory loss in the sound foot or pre-morbid sensory loss in the weak foot were excluded from the sensory testing

Testing protocol

A `Neurometer' was the only piece of equipment used. It is similar to a ballpoint pen, with the `ball point' replaced with a spring-loaded filament. The pressure required to depress the filament is used to standardise the amount of pressure applied to the subject.

Surface pressure touch: The neurometer was used to test sensitivity to pressure touch. It was applied perpendicular to the skin and pressure applied (for no more than one second) so that the filament retracted. The action of the neurometer was demonstrated on the back of the sound hand. The subject's eyes were closed. Six trials and two sham trails (when the same procedure was carried out but the neurometer was not applied to the skin) were completed for the upper and lower aspects of the MTP joint of the first toe of the weak foot. The following explanation was given to the subject: "I wish to see if you can feel this light touch on your big toe. Before each test I am going to say "do you feel this?" - don't worry if you do not feel anything in some of the trials. Remember it is important to only indicate when you feel something. Don't try to guess."

Scoring: Each successful detection scores one point with a maximum of six points for each aspect these are totalled for a maximum score of twelve. The normative data supplied in the RASP gives one incorrect detection as the impairment cut-off point (the mean score for normal subjects minus the standard deviation), so a score of less than eleven would indicate a fail for this test.

Surface localisation. This test assesses whether the subject can localise sensory input. Light pressure is applied to six different locations over the top of the weak foot and then six location over the sole of the foot using the neurometer in the way described in the pressure sensation test. The subject's eyes were closed.

The following explanation is given to the subject: " I am going to touch your foot with the neurometer. I want to show me were I have touched you."

Scoring. Each successful localisation (within 50mm) scored one point with a maximum of twelve (six for the top of the foot and six for the sole of the foot). If the subject indicated that they did not feel anything on testing then it would be repeated once. The impairment cut-off was a score of eleven.

Sensory extinction. This tests whether the subject can detect touch on the weak foot when the sound foot is also stimulated - i.e. whether bilateral sensation is detected. A different setting is used with the neurometer so that slightly greater pressure is applied. First the subject is tested with discover whether they can detect this pressure on the weak foot, if not testing is discontinued. Two neurometers are used and pressure is applied as described above to the top of the 1st MTP joint (big toe) either bilaterally (simultaneously), or singularly to the weak or sound foot. Six bilateral stimulations and two single stimulations (to each foot) are given. The following explanation is given to the subject: "You may feel me touching either you left or right foot over your big toe, or both toes at once. I want you to tell me which I am touching, say `right' , `left' or `both'."

Scoring: Only the bilateral stimulations are scored. One point is given for each successful detection with a maximum score of six. The impairment cut-off was a score of less than six, a score of 4-5 defined as mild impairment, 2-3 as moderate and 0-1 as severe impairment.

Proprioception. Two aspects of proprioception are tested: movement discrimination - whether the subject can detect movement, and direction discrimination - whether the subject can tell in which direction the joint is moving. Two joints are tested, the ankle and first toe of the weak foot. The lateral aspects of these joints are held and the joint flexed/extended passively. After one or two seconds the subject is asked whether the joint moved and in which direction (up - extension of the toe/ dorsiflexion of the ankle, or down - flexion of the toe, plantarflexion of the ankle). The subject's eyes are closed. The following explanation is given to the subject: "I am going to move your ankle/toe up and down, this is up, and this is down (demonstration given on other foot). I want you to tell me whether you can feel the joint move and whether it went up or down as soon as you feel the movement. Before each trial I am going to say `What's this?' don't worry if you don't feel all the trials, and remember to only report what you actually feel, don't try to guess."

Scoring: A point is awarded for each correct detect of movement and for each correct direction with a total score of twenty-four for the ankle and toe. An impairment cut-off of twenty-two or less is given.

Scoring of overall sensation.

All the scores are added to give a maximum score of fifty-four and an impairment cut-off of 50.

III.c.ii Neglect - The Star Cancellation Test.

The Star Cancellation test is part of the Behavioural Inattention Test. It has well demonstrated reliability, validity and sensitivity (Wilson et al 1987; Halligan et al 1989). It consists of an A4 piece of paper with large stars, small stars, letters and short words randomly positioned across the page. The subject is asked to cross out all of the small stars. There is no time limit. The paper is aligned centrally in the sagittal plane and although the subject can move his/her head during the test, s/he can not move the paper. There are 54 small stars and a impairment cut-off point of 51 stars is used.

Instructions to the subject: "This page contains stars of different sizes. This is a small star. Look at it carefully and I want you to cross out all the small stars on the page like this (demonstrate with 2 central stars). There is no time limit so let me know when you have crossed them all out."

III.c.iiiHemianopia Clinical Confrontation Test (Wade 1985)

This is tested on a pass/fail basis. The patient and tester sit facing each other, they cover opposite eyes and look at each other's uncovered eye. The assessor then moves an object (a finger or pen for example) in from the periphery of vision, keeping it equidistant form the tester and patient. The patient reports when s/he first sees the object. This point is compared with when the tester first spotted the object (assuming that the tester has a normal visual field).