MEASURING BALANCE IN THE ELDERLY:
DEVELOPMENT AND VALIDATION OF AN INSTRUMENT

A Thesis submitted to the Faculty of Graduate Studies and Research

in partial fulfillment of the requirements of the degree of

doctor of philosophy

Katherine Berg (c)

Department of Epidemiology and Biostatistics

November, 1992

ISBN 0-315-87703-0

SHORT THESIS TITLE:

VALIDATION OF A MEASURE OF BALANCE

ABSTRACT

The objective of this study was to examine the measurement properties of the Balance Scale. In two longitudinal studies, 113 elderly subjects and 70 acute stroke patients were monitored on 4 occasions during periods of one year and 12 weeks, respectively. Results showed that Balance Scale scores were associated with the occurrence of subsequent falls and clinical judgments of balance, and could discriminate subjects by their use of mobility aids and the location of their follow-up evaluation. Balance scores were strongly associated with measures of functional status and motor performance in stroke patients over the 12 weeks. Moreover, changes in the Balance Scale were able to mirror changes in the functional status of stroke patients. In addition, the Reliability Study showed excellent inter and intra-rater reliability and internal consistency of the Balance Scale when used with elderly residents and stroke patients. Overall, the results indicate that the Balance Scale has good measurement properties and is ready for use in clinical practice and research.

En Résumé

Cette étude a pour objectif d'examiner les propriétés de l'échelle d'équilibre. Deux études longitudinales ayant pour but d'évaluer deux groupes de sujets ont été répétées 4 fois pendant une période d'un an chez 113 personnes âgées, et pendant une période de 12 semaines chez 70 patients ayant subi un accident cérébro-vasculaire (ACV) aigu. Les résultats ont montré que les scores de l'échelle d'équilibre sont associés aux chutes futures et aux jugements cliniques d'équilibre, et pourraient discriminer les sujets quant à l'usage d'aides à la mobilité et le lieu où se fera le suivi. Les scores d'équilibre enregistrés durant les 12 semaines sont fortement associés aux mesures de besoins fonctionnels et de performance motrice chez les patients ayants subi un ACV. D'ailleurs, les changements de l'échelle d'équilibre ont pu refléter les changements dans les besoins fonctionnels des patients ACV. De plus, une étude de fiabilité démontre que la fiabilité et la cohérence interne de l'échelle d'équilibre sont excellentes. En général, les résultats indiquent que l'échelle d'équilibre comporte de bonnes propriétés de mesure et pourrait être utilisée en pratique clinique et en recherche.

PREFACE

Balance, as discussed in this thesis, refers to the ability to control upright posture under a variety of conditions and situations. It is not an isolated ability but rather one that is closely integrated with movement. To balance, individuals must be able to maintain certain postures, make appropriate adjustments for voluntary movements and respond to external forces. These three requirements can be considered the dimensions of functional balance. They are essential to safe mobility and the performance of daily activities.

Various biomechanical, sensory and motor impairments can adversely influence the ability to balance. In the elderly, such deficits are prevalent because of age-related changes and sequelae of injuries and diseases. The impairments are not always remediable, but neither do they necessarily lead to disability. Individuals may be taught to adapt and compensate for their impairments. One aspect of this functional adaptation involves relearning postural control within the constraints of the impairments that have occurred.

Given the close relationship between balance and safe mobility and the prevalence of impairments in this population, it is important to have a measure of balance with good psychometric properties that is appropriate for the elderly. Specifically, in the clinic, we need to evaluate the patient's ability to balance and, when indicated, use this information to plan treatments. We also need a measure of balance to monitor the course of the patient, and assess the final outcome in terms of the effectiveness of treatment. In addition, research is needed to describe the changes in balance associated with specific conditions and determine the optimal timing, content and duration of treatments. The evidence gathered from such research can be used to justify the expansion of existing programmes, to establish new services, to make better use of available resources, as well as to guide daily clinical decisions. This research also requires a quantitative and valid measure of balance.

Measurement is the procedure of applying a standard scale to a variable or set of values (Last 1988). It has also been defined as the assignment of numerals to objects or events according to rules (Stevens 1951). The broadness of both definitions suggests that the measurement of anything is theoretically possible provided there is a good set of rules to be followed (Kerlinger 1986). We are not restricted to measuring only observable physical properties such as joint range of motion, but we can assess more abstract but important concepts related to health status such as functional performance and balance.

Consequently, between 1985 and 1987, a performance-based measure of balance was developed for use as an outcome measure in research and clinical practice (Berg et al. 1989). The desired characteristics of the instrument were that it be: easy to use, portable, quantitative, comprehensive, reliable, valid and responsive to changes in clinical status in elderly subjects. The preliminary study used three different panels of patients and professionals to develop the content and examine the reliability of the Balance Scale (Berg et al. 1989). The results were promising, with good consensus about the content and excellent intra and inter-rater reliability (ICC=0.98 for each). The cross-sectional design did not permit an evaluation of all the desired properties, but did encourage the further testing of the psychometric properties of the scale.

Developing measures of abstract concepts and testing their performance is an extensive process. It is, however, important in epidemiology because measurement error can both introduce bias and adversely affect the precision and efficiency of study designs. Whereas a certain amount of error is to be expected in any measurement process, potential users should have adequate knowledge of the magnitude of expected error to assist them in judging and controlling the impact on their results (Kelsey et al. 1986).

Good measurement is essential not only in research but also in clinical practice. Use of the same measures in both areas would encourage a continuity of information between clinicians and researchers. This information can be used to describe the extent of dysfunction in various sub-groups, to assist in setting priorities for investigation, to identify where preventative and rehabilitation efforts should be focused. Given the importance of balance to independence and safe mobility, the Balance Scale could be a valuable adjunct to building this knowledge base in a variety of patient populations. Before advocating its use, further assessment of the extent to which the Balance Scale meets the standards of measurement for clinical practice and research was needed.

Therefore the objective of the present study is to investigate the measurement properties of the Balance Scale, assessing its performance in situations that closely resemble clinical reality. The information is accrued in three inter-related studies using residents of a home for the elderly and acute stroke patients admitted to general hospitals as the subjects.

The thesis is presented in five chapters. The first reviews measurement theory and the essential criteria necessary for good measurement. These criteria are considered relative to the Balance Scale: what was known from the preliminary study and what had to be addressed in the present study. This information leads to the rationale for the strategies used to accrue evidence about the measurement properties of the Balance Scale.

The second chapter is a literature review of the concept of balance, with specific emphasis on its impairment in the elderly. This chapter provides a rationale for the content of balance measures, showing the evolution of balance from an isolated sensori-motor concept to an integral part of motor performance. Various methods of assessing balance are described, but there is no one measure of balance that has demonstrated all the required properties of an outcome measure.

The third chapter presents the specific objectives and the methods used in the three inter-related studies. Validity Study I examines the concurrent and predictive criterion validity of the Balance Scale in a residence for the elderly. Validity Study II investigates the construct validity of the Balance Scale and its ability to monitor changes in the status of patients with acute stroke. The third sub-study assesses the reliability of the Balance Scale when used with the elderly residents and stroke patients.

The fourth chapter presents the results of the three inter-related studies. Each section includes a description of the participants. In Validity Study I, the results describe the association between scores on the Balance Scale and each of the criterion measures, clinicians' global judgments of balance, type of mobility aid employed and occurrence of falls during the study period. Validity Study II presents the results of the covariation among the Balance Scale scores, the Fugl-Meyer Motor Performance Scale and the Barthel Index scores for the stroke patients during the 12 weeks of follow-up. The magnitude of the change in Balance Scale scores is compared to the Barthel Index to assess if the responsiveness of the Balance Scale is equivalent to that of the Barthel

Index. In addition, the association between stroke patients' place of residence at the time of each follow-up assessment and the Balance Scale scores is presented as evidence of "known groups" concurrent criterion validity. The Reliability Study presents the results of the inter and intra-rater reproducibility of the Balance Scale as well as its internal consistency.

In chapter five the results are discussed relative to the evidence supporting its use as a clinical and research tool. The strength of the evidence of reliability and each type of validity is considered along with the generalizability of the findings. The known properties of the instrument are then summarized with regard to the characteristics of other balance measures. Lastly, the limitations of the study and the advantages and suggested uses of the Balance Scale are presented.

STATEMENT OF ORIGINALITY

The development of the Balance Scale represents original work that was begun to fulfil the requirements of a Master of Science in Rehabilitation. At the time, balance was measured clinically as global ratings or with single-item tests such as standing on one leg. In the laboratory, balance was studied in terms of postural sway, electromyography and/ or kinematic analysis. There was little consistency in assessing balance in either clinical or laboratory measurements.

The Balance Scale is the first measure of balance to have undergone a formal content development. This process was based on the judgements of professionals and the performances of geriatric subjects. Although other measures have been introduced in the intervening years, the present investigation has used a more comprehensive approach to further testing. The reliability has been assessed in situations that are found in daily practice. The criterion and construct validity have been assessed relative to a variety of external criteria and other constructs that should show a relationship to balance. The approach has been to apply psychometric theory to the development of an instrument suitable for research and clinical practice. At present, the information on properties of the Balance Scale exceeds that of the other measures and provides a substantial amount of information for potential users.

ACKNOWLEDGEMENTS

I am grateful to many individuals who have taught and guided me in pursuit of knowledge, and who have set an example of scholarship and hard work. First and foremost, I thank Sharon Wood-Dauphinee PhD, my advisor throughout my graduate studies for her leadership and her friendship. Her insight and practical grasp of critical issues in health care research are a source of inspiration. In addition, I value her technical skills as an editor.

Secondly, I would like to thank Jack Williams PhD, an ex-officio member of the thesis committee. Thank you for your guidance and expertise in the early stages of analysis. More importantly, thank you for stimulating discussions and for your comments on earlier versions of this thesis.

Thank you to the members of my thesis committee, Donna Lamping PhD, Jim Hanley PhD and Dr David Gayton. Together they offered advice and editorial assistance during the writing of the thesis. In addition, they contributed expertise in measurement, statistics and practical clinical issues. They were approachable and generous with their time.

Additionally, I am grateful to Gina Bravo PhD for her assistance in the statistical comparisons between the Barthel Index and the Balance Scale.

I am indebted to the research assistants who worked so diligently to recruit and follow the subjects: Phyllis Havercroft and Cathy Misener at the Griffith McConnell, Jacqueline Harvey at the Montreal General Hospital, and Kathryn Lexow-Nixon, Tammy Abramovitch-Ostroff, Vicki Smith, Any Brouillette and Lilianne Holtman at the Royal Victoria-MNI Complex. Thank you also to Joanne Boisvert, Monica Kilfoil, Kathleen Norman and Anna Peruzzi for their work as independent evaluators.

My thanks also extend to the staff at the Griffith-McConnell, Montreal General, Royal Victoria-MNI, Catherine Booth Hospital, Jewish Rehabilitation Hospital, Julius Richardson Hospital, Montreal Convalescent Hospital. They were interested and willing to help to facilitate the evaluations as well as participating as raters when asked. I am also grateful to all clinicians and researchers who have inquired about the Scale. This interest has been a source of encouragement.

To all the subjects and their families a special "merci". They gave of their time and let us into their homes. Their willingness to participate made the study possible.

This study was also made possible through the financial support of Health and Welfare Canada. The National Health Research and Development Program funded the project as part of the Special Competition for Outcome Measures in Habilitation and Rehabilitation.

On a more personal note, I wish to thank Lois Finch, Anita Gagnon, Binnie Gilman, Jacek Kopec, Nicol-Korner Bitensky for their last minute proof-reading of the manuscript and their helpful comments. I also owe a deep debt to Jacqueline Whitham for her assistance in making the tables.

Finally, I thank my family and friends for their support and encouragement. Words cannot fully express my gratitude.

LIST OF TABLES

LIST OF FIGURES

TABLE OF CONTENTS

Chapter I BACKGROUND AND RATIONALE

"Any strategy for altering the health status of the elderly requires a technology for first assessing that health status and then detecting increments of progress" (Kane and Kane 1981).

1.0. Introduction

This chapter reviews the basic considerations involved in judging measures - knowledge that is important for both developers of instruments and potential users. The assessment process requires consideration of the technical requirements of good measurement as well as the feasibility of using an instrument for a given purpose. Feasibility relates to how long it takes to administer the test, the ease of scoring, equipment expense, portability, training requirements of the raters, and whether the scale is acceptable to the subjects being tested. The technical questions centre on the psychometric properties of the instrument such as reliability and validity. These are discussed relative to what was demonstrated in the preliminary study and what should be examined in the present study to further investigate the usefulness of the Balance Scale in clinical practice and research.

Standards for judging measures differ in specific situations. Therefore the purposes of measurement and their necessary characteristics are briefly reviewed at the beginning of the chapter. Potentially, the Scale could be used for providing a comprehensive description of balancing ability, assessing the extent of the problem to determine a suitable intervention, monitoring the status of patients over time, and evaluating the effectiveness of treatments. These intended uses and the theory underlying instrument development provide a guide and a rationale for the strategies chosen to further assess the performance of the Balance Scale.

1.1. Purposes of measurement

In general, comprehensive measures that meet high technical standards are more likely to serve multiple purposes. However, users must balance the detail of the measure with the ease of administration and lower costs. This trade-off must be examined to assure that users do not sacrifice properties essential to their needs. Examining the properties specific to distinct purposes is also a useful exercise for instrument developers both during content development and later when testing to assess the extent to which the measure can serve multiple purposes.

Several authors have proposed guidelines for potential users with suggested characteristics for each purpose (Feinstein 1987; Kane and Kane 1981; Kirshner and Guyatt 1985). There is a considerable overlap in the classification systems. In addition, issues relating to comprehensiveness and precision of the measure vary across different purposes, depending on the particular target population and specific situation.

At any given time, measures are needed to describe the status of an individual or group. This information may be used to develop normative data on specific groups, to describe the extent of certain problems in a community or as a baseline measure for future comparisons. When selecting patients for the study, when stratifying them, or when measuring confounding variables and assessing the relationship between the variables of interest, researchers require instruments to discriminate between subjects. Use of the same descriptive measures in research and clinical practice would facilitate communication and develop a stronger knowledge base.

To be useful in any given situation, the intended instrument must have sufficient range to describe the target population and sufficient precision or detail to discriminate clinically relevant differences. In addition, the length and comprehensiveness of the instrument must be judged relative to the complexity of the concept being measured.

Screening individuals for disease or risk of an adverse event such as a fall requires a quick, efficient way of classifying subjects who are, and are not in need of further testing or intervention. The cut-off point on the scale is selected to minimize the misclassification of subjects. Therefore screening instruments need not be detailed or comprehensive.

Measuring instruments are also needed to diagnose and assess the extent of the problem and to develop an appropriate treatment plan. The amount of precision required of the instrument may differ. At times, it is only necessary to classify subjects into two groups, normal and abnormal. For other situations, it is essential to detect very small but clinically significant differences between subjects. The latter is particularly true when the assessment forms the basis for future comparison of progress or deterioration in status.

Identifying true changes in scores over time is important to monitor the status of patients and to evaluate the effectiveness of treatments. Clinicians want to know if the treatment has improved the status of their patient. Researchers in a clinical trial want to determine if the patients receiving the treatment have shown greater improvement than the control group. An outcome measure in a clinical trial must be able to detect small differences between subjects as well as within a subject over time. The greater its efficiency in demonstrating the differences, the greater the power of the study.

The last purpose to be discussed is prediction of future outcomes based on current measurements. Many decisions made in clinical practice are founded on such prognostic assumptions. For example, a patient who is discharged home is presumed to be able to manage safely in that environment.

Due to considerations of cost and time, measures generally undergo cross-sectional testing of their properties before assessing their ability to predict or to monitor changes in clinical status. At the beginning of this study, we knew that the Balance Scale was easy to administer and was portable, requiring only a stopwatch and ruler as equipment. Initial testing of the reliability and validity of the Balance Scale was promising but more extensive assessment was needed before promoting its use in clinical practice and research. Specifically, the following sections review the properties of the Balance Scale assessed in the preliminary study and consider the rationale for further testing of each property.

1.2. Psychometric properties

1.2.0. Introduction

Reliability and validity are two fundamental criteria of a good measure. These two requirements are presented in the context of how they have been addressed in the preliminary study and what still needs to be tested. In addition, responsiveness, or ability to detect clinically meaningful changes in status, is discussed as an essential criterion for an outcome measure. The rationale for further testing of each property is addressed at the end of each section.

1.2.1. Reliability

The first essential property of good measurement is reliability, also referred to as reproducibility or consistency. Measurement theory states that observed scores contain both real variation between subjects and error. Reliability is the proportion of the observed variance that is attributable to the true score differences between subjects. Clinicians must have confidence that a change in score represents a true change in the status of the patient and that score differences between subjects reflect the true diversity between patients and not variation due to observers or random error. This consistency in scoring is also very important in research as any excess variation in scores will lead to an increase in variance, which affects the required sample size and cost of a study.

To permit comparisons between the scores of different raters, instruments should be assessed for inter-rater reliability or the agreement in scoring among different raters measuring the same property. In addition, it is important to assess whether the same rater will obtain the same score at two points in time in the absence of true change. This latter quality is termed intra-rater reliability. The reproducibility of equipment or self-administered tests is termed test-retest reliability.

It is common for preliminary studies to examine reliability under controlled conditions. For example, the Balance Scale was initially assessed by having raters score the videotaped performances of 14 patients (Berg et al 1989). The results showed excellent agreement (ICC =.98) among five physical therapists, three occupational therapists and two nurses, despite minimal training in the use of the instrument. The raters also demonstrated consistency within themselves when assessing the same subjects on videotapes at least one week later (ICC =.98).

Videotaping is, however, an artificial situation that eliminates several factors occurring in clinical situations. Only one performance of each subject is videotaped. There is no possibility of inconsistencies in the performance of the subjects with repeated testing because of fatigue, improvement with practice, or lack of motivation. Similarly, videotapes eliminate any differences in the administration of the test because of imprecise instructions, inattentive scoring by the raters, or environmental distractions. Hence, it is important to repeat the assessment of inter and intra-rater reliability in less controlled conditions.

For a test in which multiple items are scored and summed, such as the Balance Scale, it is also advisable to assess how the individual items relate to each other and to the total score. This type of reliability is termed internal consistency. Each item within a scale is considered to be a single measure of the common underlying concept. Thus, summing the items provides more information and gives a more reliable estimate of the true score. In the preliminary study, the high internal consistency (Cronbach's alpha .96) indicated that the scale was measuring one underlying dimension. This characteristic facilitates the interpretation of the test results but is not essential for a measuring instrument.

The high standards achieved in the preliminary study are consistent with the levels of reliability (.98 and .95) recommended for instruments that will be used to make decisions about individuals based on repeated testing (Helmstadter 1964; Nunnally 1978). The higher standards are needed because measuring errors occur with each testing. Excessive fluctuations in the scores may mask a true change in the status of the patient and substantially alter the interpretation of the score and lead to erroneous decisions or inappropriate interventions. Group measurements in research also benefit from instruments with the highest reliability because a lack of precision will increase the variance and decrease the power of the study.

Before advocating its use, the reliability of the Balance Scale had to be assessed in situations that would allow better generalizability of the results to clinical practice. Although this further testing would incorporate more sources of potential error, estimates derived from this assessment offer more information to potential users on how to control and interpret results.

1.2.1.1. Rationale for further testing of reliability

The results of the preliminary testing showed that the Balance Scale can be scored consistently by multiple independent raters and by the same rater at two points in time. The additional sources of error that must be considered are whether different raters will administer the Balance Scale consistently, whether familiarity with the patient will influence the scoring, and whether the patients' performances are stable across different testing situations.

In the clinic, assessment of inter-rater reliability requires that the same subjects be independently evaluated by two or more raters. Similarly, for intra-rater reliability the test must be repeated by the same rater with a sufficient time interval to minimize the probability that the rater will remember the previous score but not so long as to have a true change in the subject.

To further facilitate the generalizability of the testing, raters should include the independent evaluators and health care professionals who know the patients. They should represent different professions and levels of experience and should not receive extensive training in the administration of the instrument. Patients should also reflect a range of balancing ability and have diverse characteristics to improve the generalizability of the results because it is expected that the reliablity of a measure varies according to the population being tested. Lastly, the testing should be performed in typical clinical or home environments.

1.2.2. Validity

Validity expresses the degree to which an instrument measures what it purports to measure (Last 1988). An instrument may be valid to varying degrees and valid for specific situations. When discussing validity, it is always important to consider the intended purpose of the measurement.

Measurement theory defines validity as the proportion of the variance of a measure that is shared by two or more tests. This common factor variance contains neither error variance nor variance that is specific only to that one instrument (Kerlinger 1986). Reliability therefore is a necessary but not sufficient condition for validity. Scores may show consistency and still not be measuring the intended concept or attribute.

While the measurement of any attribute is theoretically possible, there is greater complexity involved in the measurement of a behaviour or an abstract concept than measuring a physical property. In particular, the challenge lies in demonstrating its validity because this quality must be considered relative to a particular purpose. Validation of an instrument requires the use of several strategies to examine whether the scale is performing to expectation. The three major types of validity are: content, criterion and construct.

Content validity refers to whether the items in the scale adequately represent the dimensions and domain of the concept of interest. The decision is basically a judgmental one, but the plan and procedures of instrument construction help to assure its validity (Nunnally 1978). Defining the domains and dimensions of the concept begins with a thorough search of the literature. The process may also include a systematic questioning of experts or individuals working in the area (Spitzer et al. 1981; Wood-Dauphinee et al. 1987). The large pool of potential items is gradually narrowed to produce a scale that is sufficiently comprehensive and of an appropriate length to be practical.

The content of the Balance Scale was developed in three phases, with a different panel of geriatric patients and health care professionals participating in each stage. The elderly participants had either fallen recently or were receiving physical or occupational therapy treatments. Following assessment of internal consistency, two questionable items were dropped, leaving 14 items each scored on a five point scale. The remaining items are movements or tasks common to everyday life such as standing up from a chair, turning, and picking up an object from the floor. The scoring is based on an independence-dependence continuum in addition to specific time and distance requirements. A copy of the Balance Scale is included in Appendix 1.

At this stage, there was reasonable evidence for the content validity of the Balance Scale for geriatric subjects. Beyond content development, it was necessary to further demonstrate the extent to which this scale measured balance. In the absence of a gold standard for assessing balance, several different strategies had to be considered. While there is some controversy on the names given to these types of validity, the general objective remains the same: to assess the degree to which an instrument performs relative to other measures or in situations that are consistent with theoretical expectations. The nomenclature used in this thesis is outlined in the following paragraphs.

The first major category, criterion validity, can be sub-divided into two types: concurrent and predictive. The most convincing evidence of the validity of a new instrument would be to show a strong correlation between the scores on that instrument and scores on an existing gold standard, a universally accepted valid measure. If the two measures are obtained at essentially the same time, it is called concurrent criterion validity. Given the rarity of a gold standard, correlations are often made against other existing measures. These correlations are limited by the validity of the comparison measure but do comprise an acceptable and realistic approach to assessing validity.

Another way of assessing concurrent criterion validity is to choose a criterion variable that identifies different levels of status or impairment that are relevant to the construct being assessed. The distinguishing feature of this type of criterion validity is that the external criterion is categorical rather than continuous (Spector 1992). This method of assessing validity is also referred to as "known groups" technique (Bohrnstedt 1983; Spector 1992).

The second type of criterion validity examines scores on a scale in relation to the occurrence of a future event. It is called predictive criterion validity and is an important property for health status assessments because many decisions in clinical practice are based on prognostic assumptions.

Whereas all validation procedures involve the assessment of the instrument's performance, construct validation is specifically concerned with testing the theoretical framework underlying the scale (Carmines and Zeller 1979; Kerlinger 1986; Nunnally 1978). One aspect of construct validity addresses the meaning of the test scores and what factor or factors can account for the variance of the test scores. Factor analysis and item-to-total scores can assist in assessing the relationships between items in a scale and what proportion of the variance can be explained by different factors. The other approach to construct validity involves generating hypotheses about how the construct relates to other constructs and assessing the degree to which its performance is consistent with expectations stated apriori. For example, it is reasonable to expect that scores on a balance measure should demonstrate an association with related constructs such as mobility, functional status, and motor performance.

1.2.2.1. Rationale for further testing of validity

Criterion validity

Often, the first comparison for a new scale examines its relationship with clinical judgments. In the clinical setting, an individual's ability to balance is commonly described as good, fair or poor. Such global ratings are prone to subjectivity as each rater uses his or her own reference point. Nonetheless, these judgments are used in practice, and should demonstrate a relationship with scores on the Balance Scale, providing evidence of concurrent criterion validity.

Another criterion associated with balance is the use of mobility aids such as canes or walkers. A walker is an aluminum frame that is used to reduce the amount of weightbearing through the lower extremity by allowing some weight to be distributed through the arms. Walkers are also used to compensate for impaired balance by enlarging the base of support. They are most often used indoors, but can have wheels and/or be collapsible for easy storage and transportation. Canes offer less support than a walker but do compensate for minor postural instability. It is not uncommon for individuals to use a cane only outdoors due to the greater uncertainty of environmental conditions and the fear of being jostled by a crowd. On average, elderly individuals who walk without an aid may be assumed to have better ability. Physical capability is not the only determinant of the use of a walking aid. The decision may be influenced by other factors such as cognitive impairment, poor judgment, or vanity. Nonetheless, a Balance Scale that is able to discriminate among subjects by their use of mobility aids would demonstrate the "known groups" type of concurrent criterion validity.

In the early recovery period, stroke patients remain in general hospitals, are sent to a rehabilitation centre or are discharged home. The criteria for discharge vary according to the availability of community support but, on average, those discharged home have a higher level of functioning and are considered to be able to manage their daily activities safely. Patients transferred to a rehabilitation setting remain for a two to three month period. The timing of the transfer is influenced by the availability of beds, the resolution of any acute medical problems, and is contingent on the patient demonstrating a potential for discharge. The remaining patients undergo treatment in the general hospital. Some spend years in extended-care wards awaiting placement in nursing homes. Hence, the location of each follow-up evaluation of stroke patients is an indicator of differing levels of ability and can also be used to assess the "known groups" type of concurrent criterion validity of the Balance Scale.

The most pertinent criterion relative to the elderly population is the occurrence of falls. Whereas it need not predict all falls, it is logical to expect a relationship between balance and the future occurrence of falls. To assess this relationship it is helpful to have subjects who are independently mobile, have a wide range of ability and live in a defined area such as residential care. The latter requirement facilitates recruitment and evaluation of subjects, and improves the probability of accurate reporting of falls.

Construct validity

Patients with stroke often display impairment in balance, motor performance and their ability to carry out basic activities of daily living. Although not all patients will improve, those who do demonstrate the greatest gains in the first few weeks. The expected changes in performance make this population an ideal group to assess both validity and responsiveness to changes in status.

When considering the balancing ability of stroke patients with motor impairment, it can be hypothesized that any changes in balance would parallel changes in motor performance and function. The interrelationships between balance, motor performance and functional status provide a rationale for expecting covariation between these measures over time. The feasibility of this strategy for construct validity is facilitated by the availability of well developed scales that assess motor and functional performance.

1.2.3. Responsiveness

Responsiveness has been defined as the ability of an instrument to detect clinically important changes in the status of the subjects, even if they are of a small magnitude. The property is considered essential to measures used to evaluate the effectiveness of interventions (Guyatt et al. 1987). When considered as a requirement for an outcome measure, responsiveness also incorporates the ability to discriminate small but meaningful differences among subjects. For example, in a clinical trial, an outcome measure must be able to detect improvements in the status of the treated patients that are greater than the changes in the control group. An instrument that can more efficiently detect small changes and discriminate small clinically relevant differences among subjects is a more responsive instrument and a better choice as an outcome measure (Kazis et al. 1989; Liang et al. 1990; Norman 1989; Tuley et al. 1991).

The responsiveness of several outcome measures in a study can be compared using effect size, a standardized measure of group differences or responses to change (Cohen 1977; Kazis et al. 1989; Liang et al. 1990). The instrument with a large effect size is more efficient in terms of the power of the study and sample size requirements than one with a smaller effect size. Variations exist in the computation of the effect size. To adequately address the variability of response to treatment, it is preferred to use a formula that uses the standard deviation of differences or changes rather than a baseline or pre-treatment standard deviation as the denominator (Liang et al. 1990). Similarly, the effect size is preferred to the Index of Responsiveness suggested by Guyatt and associates (1987) because the latter only considers the amount of clinically meaningful change relative to the variability in a stable control group and therefore, is likely to underestimate the variability of scores in a clinical trial (Norman 1989; Tuley et al. 1991).

Other methods of assessing responsiveness include a statistic based on the t-test that examines the relative efficiency of any two measures in evaluating a treatment of known efficacy (Liang et al. 1985). New scales have also been compared to existing physiological and clinical tests known to change with a treatment of known efficacy (Meenan et al. 1984). Additionally, the sensitivity and specificity of identifying true change, an analogy with diagnostic tests, has been used (Deyo and Centor 1986; Deyo and Innui 1984; MacKenzie et al. 1986).

The above approaches are unrealistic for the present study for several reasons. Firstly, the present study has no planned intervention. Most strategies require an intervention, preferably of known efficacy. The use of sensitivity and specificity presupposes that a change can be identified with certainty. In the absence of a gold standard and a treatment of known efficacy, the decision appears subjective and unreliable. This method also requires a dichotomous outcome, change or no change, resulting in a loss of information for a continuous variable such as the Balance Scale.

1.2.3.1. Rationale for assessing responsiveness

Stroke patients can undergo marked changes in status during the early recovery period. The present study examines the changes in balance occurring in stroke patients at 2, 4, 6 and 12 weeks post onset. The strategy for assessing responsiveness first addresses whether the Balance Scale can detect changes in the status of the acute stroke patients over the 12 week follow-up period. The next step is to compare the magnitude of the changes in the Balance Scale scores to an external criterion recognized as an outcome measure in stroke rehabiliation.

The Barthel Index (Mahoney and Barthel 1965) is chosen as the criterion measure. It assesses self-care and basic mobility skills, factors that are expected to show marked improvement in the early recovery period. Moreover, it has good measurement properties and is widely used in stroke outcome research (Chino 1990; Hewer 1990; Granger et al. 1988; 1989; Reding 1990). Wade and Collin (1988) advocate its use as a standard index in clinical practice and research. In addition, the Barthel Index has been shown to be more efficient in demonstrating a treatment effect in acute stroke patients as compared to the Fugl-Meyer Motor Performance Scale (Fugl-Meyer et al. 1975) and measures of neurological status and stroke severity (Wood-Dauphinee et al. 1990). A copy of the Barthel Index is included in Appendix 2 and the properties are further discussed in later chapters.

1.3. Summary

In summary, the Balance Scale is a performance based measure that is portable, takes 10-15 minutes to administer and is feasible for use in a wide variety of testing situations. Its measurement properties showed promise in preliminary testing but required further investigation to assess the usefulness of the measure as a clinical and research tool. In particular, use of the Balance Scale to monitor the status of patients and evaluate the effectiveness of treatments requires that the measure detect clinically meaningful changes and meet high standards of reliability.

Measurement theory describes several types of validity and how they are assessed. Each can be used to accrue evidence for the validity of the Balance Scale using criteria relevant to balance such as clinical judgments of balancing ability, occurrence of falls, use of mobility aids, and location of follow-up evaluations. In addition, the expected covariation between balance, motor capacity, and functional performance in stroke patients during the early recovery period provides a theoretical basis for evaluating construct validity. It also allows an estimate of the ability of the scale to monitor changes in the status of the stroke patients relative to the Barthel Index, an instrument commonly used as an outcome measure in this population.

The merits of a measure will become better known as it becomes more widely used. The test developers must, however, accumulate a certain body of information before encouraging the use of a new measure. The present study addresses these issues in examining the Balance Scale.

Chapter 2 LITERATURE REVIEW

2.0. Overview of the chapter

When studying balance and its measure, it is important to understand the intended concept and to have knowledge of alternative tests or instruments. The first part of the chapter explores the concept of balance and considers the requirements for normal postural control. This knowledge provides a rationale for measuring balance. The second section will review the changes that interfere with postural control and make impairment of balance a source of concern in the elderly. Lastly, the measurement properties of existing balance tests will be discussed. This information provides further justification for the continued testing of the Balance Scale.

2.1 Concept of Balance

Balance in its broadest sense includes the capability to control upright posture under a variety of conditions and situations and the ability of an individual to sense his or her limitations of stability. This definition allows for the close interaction of movement and postural control, and represents a progression from the earlier models of motor control that considered balance to be an isolated sensorimotor concept.

The early work on the function of the sensory systems was performed by Sherrington (1947) with specific contributions relative to posture by Magnus (1926). By stimulating specific sensory systems in decorticated animals, they elicited stereotyped responses or reflexes. This research became the basis for a model of motor control that believed that reflexes were fundamental to all human movement. Under this model, the basis of postural control was believed to be a chain of reflexes which contributed to maintaining posture and orienting the body as dictated by situations from the outside world. Hence, sensory inputs were considered essential for motor outputs.

This reflex hierarchical model of motor control has been criticized on the basis of studies demonstrating that reflexes can be modified by learning, that coordinated movement can occur in the absence of intact sensory systems, and that movements can be initiated before any sensory stimuli (Belenkii et al. 1967; Horak 1991; Polit and Bizzi 1979; Thelen et al. 1987).

Currently, postural control, movement, coordination and skill development are being studied under the more global sphere of motor control. More recent models of motor control have been proposed (Bernstein 1967; Horak and Nashner 1986; Nashner and McCollum 1985; Reed 1982; Thelen et al. 1987) to allow for greater flexibility in studying the complex interactions of various neural, mechanical, and behavioral factors that influence motor development and postural control.

Reed (1989) proposed that future studies should be directed at assessing how humans perform tasks under a variety of functional contexts rather than how they respond to artificial situations. This recommendation is compatible with clinicians' judgments on the necessary components of a balance scale (Berg et al. 1989) and the trend toward performance-oriented tests in the evaluation of the elderly (Canadian Task Force on the Periodic Health Exam 1979). Given the ultimate goal of improving the functional status of the client, it appears most useful to define the dimensions of the concept of balance as three basic functional requirements: maintaining a given position, remaining stable while moving voluntarily and being able to adjust to external disturbances or irregularities in the environment. Ideas for assessing each dimension can be gained from examining how balance has been studied in a variety of disciplines.

2.2. Rationale for the content of balance measures

Normal postural control involves several factors; impairments in any one area may adversely affect the efficiency of postural adjustments. To better understand the requirements of normal postural control it is useful to review how balance has been studied in terms of biomechanics, motor responses, sensory systems and the relationship of voluntary movements and postural control.

2.2.1. Biomechanical considerations

Maintaining upright posture can be considered in terms of simple biomechanics. A stationary body is most stable when a vertical projection from its centre of gravity falls in the middle of its base of support. Greater relative stability is achieved by increasing the area of the base of support or lowering the centre of gravity; hence, sitting is more stable than standing. Similarly, activities such as reaching forward or putting on a sweater are easier while sitting than standing because the larger base of support safely permits greater displacement of the vertical projection of the centre of gravity before it reaches a critical boundary at the periphery of the base of support. Once such a critical point is reached without appropriate corrective action, a loss of balance or a fall would occur.

When standing with an ideal "normal" posture, a vertical line, drawn through the body's centre of mass at the level of the second sacral vertebra, can be viewed from the side to fall 5 centimetres in front of the ankle joint, just in front of the centre of the knee joint, through or just posterior to the hip joint, in front of the shoulder joint, and through the mastoid process (Galley and Forster 1987). Such an ideal body posture reduces the amount of muscle work required to stand still (Basmajian and Deluca 1985). However, there is always a subtle amount of movement or postural sway due to the alternating action of antagonistic muscle groups and the effects of gravity, resulting in continuous minor weight shifts and changes in body alignment. Shifts also occur from side to side, alternating the main support from one leg to another. This constant shifting of the body's centre of gravity helps prevent fatigue and assists in venous return (Galley and Forster 1987).

The amount of postural sway exhibited by an individual has been used to indicate stability while maintaining a standing position or in response to the movement of a supporting platform. It is assumed that persons who sway less are more stable because the projection of the centre of gravity stays closer to the middle of the base of support (Hasselkus and Shambes 1975). Similarly, the less the postural sway in response to an external disturbance, the more stable is the individual (Maki et al. 1990).

Clinical tests commonly incorporate the principles of biomechanics to assess the ability to maintain positions of increasing difficulty by altering the base of support. For example, subjects may be timed while standing on one leg or with one foot directly in front of the other.

2.2.2. Pattern and sequencing of motor responses

A necessary component of balance is the ability to make the appropriate postural adjustments or responses. Muscle weakness or paralysis can severely limit this capability. Performance-based measures indirectly assess strength when they evaluate the degree to which the task was accomplished and the insecurity associated with the movement. Balancing tasks perceived as being more difficult often also require greater strength. For example, ankle evertors and invertors must work harder to offset the biomechanical disadvantage of standing cn one leg instead of two. In addition to the ability to generate and maintain sufficient muscular force, effective postural adjustments require muscles to respond in a sequence and pattern that is efficient for the task at hand.

Nashner and colleagues (1976; 1977; 1979) developed an apparatus able to give both translational and rotational perturbations. Using electromyography, they studied the onset and sequencing of muscle responses of subjects following such perturbations. They identified two primary strategies. The most common pattern of response to the platform movements, termed the ankle strategy, involves shifts of the centre of body mass as a rotation about the ankle joint with little or no movement of the hips (Nashner et al. 1977;1979;1987). The other response sequence, the hip strategy, shifts the centre of body mass by flexing or extending at the hips.

Nashner and colleagues (1985) later indicated that subjects may mix strategies and that certain subjects may bend their knees for greater stability, a tactic referred to as the suspensory strategy. Detection of this alternative strategy requires kinematic information to be collected in conjunction with EMG and force plate analysis (Patla et al. 1990). Additionally, when the perturbation has been too fast or large, or if the ankle or hip strategy could not be used in sufficient time, the subject may take a step to keep or bring the centre of gravity within the base of support (Horak 1987; Shumway-Cook and Horak 1986).

The size of the perturbation, the type of support surface, and length of the support surface are other factors that have been linked to a particular pattern of response. While there may be a limited number of response synergies, it is not clear that all normal subjects respond to the same stimuli in the same way. A limiting factor to the generalizability of the findings of Nashner (1976; 1977) was that all study participants had to demonstrate a strong response of the ankle strategy. Individuals excluded from the study may have demonstrated greater variability in response patterns.

Further research is needed to assess the acceptable range of variability in postural responses in different functional contexts. This information may permit earlier identification of impairment and lead to more appropriate intervention strategies.

2.2.3. Influences of the three sensory systems on postural control

Balance has also been studied in terms of the relative contribution of each of the three sensory systems involved in postural control: somatosensory, visual and vestibular. Information concerning the movement of body segments with reference to each other and to the support surface is given by the somatosensory system. The contributions of this system have been studied by asking subjects to stand on surfaces that do not provide accurate somatosensory feedback. For example, they have been asked to stand on thick foam (Horak 1987; Shumway-Cook and Horak 1986) or on a platform that is oscillating at the same frequency as the postural sway of an individual (Nashner et al. 1982).

The visual system provides information about the body's position relative to the environment. It also works in combination with the vestibular system to stabilize gaze when movement occurs, a function important for spatial orientation. Individuals must be able to stabilize their gaze while moving their head or body, and when there is a movement of a large portion of the visual field, such as passing traffic. When functionally appropriate, the vestibular-ocular and optokinetic responses can be suppressed to allow the eyes to fixate and track moving objects (Leibowitz and Shupert 1985).

Visual inputs can be eliminated by asking subjects to close their eyes or wear a blindfold (Berg et al. 1989; Dornan et al. 1978; Maki et al. 1987; 1988). The effects of inaccurate visual inputs may be studied using enclosures or domes that move at the same frequency as the postural sway of the subject (Horak 1987; Nashner et al. 1982; Shumway-Cook and Horak 1986; Woollacott et al. 1982). The rationale for testing both the absence of visual input and inaccurate information is that postural control may be differentially affected by the two situations.

Under usual sensory conditions, vision and somatosensory information dominate the control of orientation and balance. However, it is the vestibular system that is essential for resolving conflicting sensory information. (Black et al. 1983; Nashner et al. 1982). It provides input concerning the position of the head in relation to gravity, and during movement, the linear and angular acceleration of the head. The role of the vestibular system has been studied in individuals with known vestibular deficiencies (Black et al. 1983) and in experiments that provide inaccurate visual and somatosensory inputs to normal subjects (Nashner 1971; Woollacott et al. 1986).

2.2.4. Relationship between voluntary movement and balance

Basically, any movement of the body will potentially cause a shift in the centre of gravity, but this movement is usually accompanied or preceded by the counterbalancing of another body segment so that the body's centre of gravity stays close to the centre of the base. The functional significance of this shift may be to minimize energy expenditure, to increase the mechanical efficiency of the main movement or to prevent an actual loss of balance due to the movement of the whole body or a major body part (Gahery 1987). In fact, postural adjustments occur in advance of, during and following a voluntary movement (Gahery 1987).

The close association beween postural adjustments and voluntary movement has been studied using electromyography, force plates and kinematics (Belenkii et al. 1967; Cordo and Nashner 1982; Hayes and Riach 1989; Horak et al. 1984). The postural responses occurring prior to the movement can be predicted based on the starting position, the trajectory, and the velocity of the movements (Bouisset and Zattara 1987; 1981; Hayes and Riach 1989). However, if external support is given, the responses are diminished or extinguished (Cordo and Nashner 1982).

Given the importance of this relationship, one important dimension of balance is remaining stable while moving voluntarily. Therefore, balance can be evaluated by observing the ease and safety with which subjects perform movements common to everyday life: standing up from sitting, turning, and picking up an object from the floor.

Ease and safety can be scored relative to time requirements or to the degree of supervision or assistance needed for the subject to complete the movement. This type of assessment is not only functionally relevant, it is consistent with current theories and practice. Because the efficiency of these postural adjustments are influenced by biomechanical, motor and sensory changes, these factors are indirectly evaluated in the assessment.

Summary

Postural control has been studied by biomechanists, kinesiologists, neurophysiologists and others. They have examined subjects' ability to maintain upright postures, remain stable while moving and react to external forces. No one discipline offers a comprehensive approach to measuring the concept of balance, but together, the studies have helped define the domains of balance. Moreover, the close links found between postural control and movement provide a rationale for clinical evaluations of balance based on the performance of movements.

2.3. Balance in the elderly

2.3.1. Falling, a functional problem of balance

Balance is of special concern in the elderly due to its relationship with mobility and functional independence. Given an aging population and the rising costs of medical care, it is important to examine methods of preventing or slowing the decline of postural control in the elderly and treating dysfunction associated with diseases, such as stroke, which are common in this age group.

A substantial amount of research has been done to identify the risk factors for falls in the elderly and to screen for individuals who require intervention. It is difficult to compare the results across the studies due to differences in the populations, methods and measurement procedures. Nonetheless, impaired balance has been identified as a risk factor for falling, both in prospective studies (Campbell 1989; Tinetti et al. 1986; Tinetti et al. 1988) and in cross-sectional studies comparing elderly individuals with a history of falls to those who have not fallen (Chandler et al. 1990; Lipshitz et al. 1991;

Wolfson et al. 1986).

Consistency of findings is present despite differences in the operational definitions of balance which range from performance-based tests (Tinetti et al. 1986; Tinetti et al. 1988) to the ability to withstand a mechanical force pulling backward at the hips (Chandler et al. 1990; Wolfson et al. 1986) to postural sway (Fernie et al. 1982; Gabell et al. 1985). Other studies did not measure balance but included factors that can be considered as proxies or variables closely related to balance. Examples include lower extremity dysfunction, impaired gait, and use of mobility aids (Campbell et al. 1990; Grisso et al. 1991; Morse et al. 1987; Prudham and Grimley Evans 1981). Additionally, in the absence of a direct measure of balance, diagnoses such as stroke, arthritis and Parkinsons's disease have been identified as risk factors for falling (Grisso et al. 1991; Mayo et al. 1989).

There are many factors, other than balance, involved in falls. Certain individuals may learn to compensate for deficits while others push themselves to the limits of their capability or take greater risks relative to their level of ability. To date, there is no good way of evaluating the amount of risk an individual takes within his or her daily activities.

Regardless of other predisposing factors, a fall results from a loss of balance. Technically, the vertical projection from the centre of mass of the body reaches a critical point at the limits of the base of support without a corresponding postural adjustment. Avoiding a fall requires a timely and coordinated response of appropriate magnitude as well as an accurate self perception of the limits of stability.

Falls do not always result in a major incident. Errors in movement or postural adjustments are expected when acquiring new skills. Young children frequently fall as they explore the environment and learn the limits of their ability. Adults, engaging in sports such as skiing, voluntarily displace their centre of mass closer to the edges of the base of support; thus, often approaching the critical limits of stability and sometimes falling. Falls become worrisome when they occur during activities that were previously performed safely. Individuals who feel vulnerable and unsteady may compensate by restricting their activities. This restriction may lead to greater impairment through disuse, and ultimately to functional dependence.

Any disease, injury, or age-related change that affects the coordination, timing or force of the muscular responses or the ability to utilize sensory information will adversely affect balance and may predispose the individual to falling. The evidence presented in this review comes from cross sectional studies that have compared elderly persons with or without a history of falls, and young and old subjects, as well as a few longitudinal studies of older adults. It is not clear whether the documented changes occur as a direct result of aging, sub-clinical manifestations of diseases, or disuse related changes. Irrespective of cause, there are differences in the elderly that contribute to their vulnerability to falls. The information is grouped under similar headings to the prior review of the ways in which balance has been studied: biomechanical, muscle responses, sensory systems and other factors. Impairment in any of these areas have the potential to affect the efficiency of postural responses.

2.3.2. Factors influencing balance in the elderly

2.3.2.1. Biomechanical considerations in the elderly

Deviations from the ideal normal postural alignment are recognized to occur with advancing age. One or more of the following changes in the axial skeleton are believed to be present in most persons over the age of 60: a head forward position, a thoracic kyphosis, and flattening of the lumbar spine (Kauffman 1990). Muscular contractures and diminished joint range of motion may occur from changes in the connective tissue of aging muscle and the cross-linking of collagen (Kauffman 1990) or from acquired leg length discrepancies or joint degenerative changes. Additionally, arthritic conditions, among others, are associated with joint pain and/or swelling, factors which may have an independent effect on postural alignment.

As a result of such biomechanical changes, the optimal length or angle of pull of the postural muscles may be compromised. Additionally, the relative position of the vertical projection of the centre of gravity may be altered, decreasing the functional area of the base of support. This situation may require greater or more frequent postural adjustments instead of the subtle agonist/ antagonist contractions associated with postural sway in younger or more fit subjects.

Irrespective of cause, studies have shown that the elderly tend to have greater amplitudes of postural sway than younger adults (Overstall et al. 1977; Sheldon 1963;). At least one biomechanical factor, presence of a scoliosis or curvature of the spine has been associated with increased postural sway (Brocklehurst et al. 1982).

Retraining balance involves both treatment of the underlying biomechanical constraints and adaptive training under a variety of conditions. There is evidence that flexibility and range of motion can be improved in the elderly who undertake exercise programmes (Hopkins et al. 1990; Morey et al. 1991; Raab et al. 1988). Other changes in body alignment may be irreversible but need not lead to disability. Rehabilitation and appropriately structured practice may play a role in teaching individuals to adapt to their new body configuration or to learn compensatory strategies. Without this additional training, the level of activity of older individuals may be curtailed because they cannot adjust to their own age-related changes.

2.3.2.2. Differences in motor responses among the elderly

There are indications that deficits in strength contribute to falling in the elderly. Three cross sectional studies found that persons with a history of falls had diminished lower extremity strength when compared with similar elderly subjects who had not fallen (Gehlsen and Whaley 1990; Studenski et al. 1991; Whipple et al. 1987).

Assessment of the association between balance and strength is difficult to quantify due to the differing methods of assessment within the studies. Gehlsen and Whaley (1990) operationally defined balance as the ability to stand on one leg with eyes open and closed and the number of missteps while walking backward on an eight foot line. Studenski et al. (1991) assessed balance by asking subjects to stand up, turn around, stand on one leg and react to a backward perturbation at the hips (Wolfson et al. 1986). Despite the differences in measuring procedures, the findings are in agreement and point to the need for studies to examine an intervention of both strength and balance in an elderly population.

In all three studies, the fact that the strength measurements were taken after the fall, makes it difficult to clearly establish whether strength loss occurred prior or subsequent to the fall. In some cases, the testing was performed many months after a fall (Gehlsen and Whaley 1990; Studenski et al. 1991; Whipple et al. 1987). Subjects may have limited their activity level to lower their risk of another fall. This decrease in physical activity could have contributed to a loss of strength post fall.

Evidence for musculoskeletal impairment preceding disability in the elderly comes from a longitudinal population based study (Jette et al. 1990). The Massachusetts Health Care Panel Study surveyed 1,625 persons aged 65 and over on four occasions over a 10-year period but included an evaluation of physical impairment only at six and ten years (Jette et al. 1990). Impairment was assessed as subjects performed 10 body movements. The degree of disability was represented as dependency in basic and instrumental activities of daily living (ADL) items and difficulty in selected physical performance items.

The results demonstrated that progression of impairment was related to increasing difficulty in physical performance and disability. Moreover, the type of disability was related to specific impairment. For example, hand impairment was associated with disability in basic ADL functions such as dressing and feeding; whereas lower extremity impairment was linked to deficits in instrumental ADL, which included shopping and banking. Lower extremity impairment was most common in women and participants aged 80 and over.

The reason for diminished strength in the elderly is likely due to anatomic and physiological changes as well as a reduction in physical activity. Age-related changes include a decrease in the number of muscle fibres, a reduction in total muscle mass, and a diminished cross sectional area of Type II muscle fibres (Grimby 1990). This loss of Type II fibres is not uniform throughout the body, leading to the hypothesis that altered physical activation patterns may contribute to the observed changes.

Whether occurring prior to, during, or in response to a perturbation, efficient postural adjustment requires muscular responses appropriate to the situation. Any changes in the elderly that influence the timing and onset or the response or the strength of the muscles may adversely affect balance.

Woollacott and associates (1982;1986) compared the sequence and timing of motor responses of young and old healthy subjects. The elderly subjects were more prone to demonstrate antagonist muscle activation, more variable in the amplitudes and sequencing of muscle responses within each synergic grouping, and more likely to demonstrate a hip strategy than younger subjects in the same situation (Woollacott 1982; 1990). Older subjects also demonstrated a prolonged latency for the onset of the postural response particularly in the tibialis anterior muscle (Woollacott 1990; Manchester et al. 1989).

The co-contraction of agonist and antagonist could be explained as a stiffening strategy that makes it easier to maintain balance by reducing the movement options of the musculoskeletal system (Woollacott 1990). Stiffening has been observed during laboratory tests of sway (Maki et al. 1991) and during the acquisition phase of learning a new motor skill (Higgins 1991).

Inglin and Woollacott (1988) studied the anticipatory activation of postural muscles in advance of reaction-time arm movements. They found that older subjects demonstrated longer latencies in postural and voluntary muscle activation when asked to push or pull a handle. The authors hypothesized that either the voluntary control system slows with age, limiting the speed of voluntary movement, or the deterioration of the postural control system limits the speed of the movement.

The pattern and sequencing of the postural responses have been found to be so variable in the elderly that they cannot be a good indicator of the degree of impairment. Despite discernible differences in clinical tests of balance, strength and range of motion, Studenski and colleagues (1991) were unable to differentiate between fallers and non-fallers on the basis of postural response strategy. Additionally, Duncan and associates (1990) found different muscle response latencies and patterns elicited in the same subjects with changes in the type of perturbation.

Clearly, the pattern and sequencing of responses differ across individuals and within the same individual in different contexts. More information is required to assess when individuals exceed the range of normal variability, and how to retrain them to have more efficient and safer responses.

2.3.2.3. Sensory contributions to loss of balance

There is anatomical and physiologic evidence of changes in the three sensory systems in the elderly. Horak and colleagues (1989) carefully reviewed the common organ pathologies affecting the sensory systems of this age group. The somatosensory system shows losses in vibration, cutaneous and joint position sense in the ankles (Potvin et al. 1980; Stelmach and Worrington 1985; Wanger and Wang 1974). Anatomical studies of the vestibular system have noted degenerative changes with age (Horak et al. 1989; Rosenhall 1973; Rosenhall and Rubin 1975). Age-related losses in visual acuity, depth perception, contrast sensitivity, and impaired "pursuit eye movements" have also been documented (Cohn and Lasley 1985; Horak et al. 1989). Opthalmalogic diagnoses frequently found in the elderly include chronic glaucoma, cataracts, and macular degeneration.

Visual disturbances are particulary important when observing a moving object or during dynamic postural stabilization (Isaacs 1985; Woollacott et al. 1982). For example, elderly subjects may have difficulty orienting themselves with the environment if they cannot perceive slowly moving objects or if passing images persist too long. These factors may explain why visual impairment, in one or both eyes, has been identified as a risk factor for hip fracture secondary to a fall (Felson et al. 1989; Grisso et al. 1991).

Manchester and colleagues (1989) studied the differences in the pattern and sequencing of motor responses between young and old subjects under conditions of altered sensory inputs using an experimental paradigm similar to previous studies (Woollacott et al. 1982; Nashner et al. 1982). The somatosensory inputs were manipulated to provide inappropriate information for maintaining balance by setting the frequency of platform sway to correspond to the subject's own sway, and by keeping the ankles at a 90 degree angle. Visual information was not altered by an enclosure. Instead, the visual field was manipulated with the use of different types of goggles; translucent, peripheral vision occluded and foveal vision occluded. These visual conditions simulate the problems encountered by individuals with visual impairments such as cataracts and glaucoma.

The results demonstrated differences in muscle sequencing and agonist-antagonist co-contraction. Additionally, losses of balance occurred when both somatosensory and vision were simultaneously giving inaccurate information. Most commonly, the elderly subjects lost their balance when their peripheral vision was occluded and if some evidence of underlying pathology was found in the clinical screening tests.

Warren and associates (1989) noted age differences in the ability to perceive the direction of self-motion from optical flow. They did not find evidence for differences in strategies but rather a general decline in the ability to detect and localize global optical flow patterns, which may have implications for high speed locomotion and falls.

Any degeneration or pathology present in the sensory systems can potentially interfere with efficient postural control. Fortunately, under normal circumstances there is redundant information available. Wc are aware of the need to assess the impact of a major disease or injury; however, to date little is known about the cumulative effects of minor multisensory deficits associated with a single diagnosis such as diabetes.

2.3.2.4. Other factors related to balance

The capacity to adapt to changing internal or external situations is essential to balancing ability. Medications, emotional factors and cognitive ability can adversely influence judgment, alertness and performance.

Cognitive changes can occur in the elderly and have been linked to increased incidence of falls (Tinetti et al. 1988; Vlahov et al. 1990). Depending on the severity, cognitive changes may drastically affect the functional status of the individual or cause a deterioration in the performance of complex tasks. Specifically related to balance, cognitive impairment may cloud judgment about safe behaviors. It may also be a limiting factor in the ability to compensate for a deficit or to relearn motor skills, that have deteriorated through injury or disease.

Depression, anxiety and fear are other emotional variables that may indirectly influence activity level and motor performance. Similarly, distracting factors such as pain, being startled or conversations may hinder an appropriate response. Fear of falling (Maki et al. 1991) and inattention (Stelmach et al. 1990) have been associated with poorer performance in balance tests.

Finally, certain medications, in particular psychotrophic drugs, have been associated with a greater risk of falling in the elderly (Grisso et al. 1991; Ray and Griffin 1990; Tinetti et al. 1988). There may be several mechanisms explaining the effect of medications on the incidence of falling. For example, sedatives and tranquillizers, particularly those that remain in the blood stream for long periods, may decrease awareness of potential hazards or slow the individual's motor response. Haldol, a phenothiazine and a very potent tranquillizer, has motor side effects similar to Parkinson's Disease, and has been associated with the risk of falling. Patients taking medications for pain, hypotension or cardiac disease may demonstrate an increased risk of falling because the underlying pathology predisposes them to falls. Additionally, the total number of drugs is considered a risk factor for falls, as it likely represents a proxy for the poorer health status of fallers relative to non-fallers (Lipshitz et al. 1991; Tinetti et al. 1986).

Summary

Falls are of concern in the elderly due to the risk of serious injury and the possibility of an associated decline in physical functioning. At present it is difficult to compare across studies due to differences in subjects, methods and measurements. Balance has been consistently identified as a risk factor for falls. In addition, there is evidence for biomechanical, motor, sensory and cognitive changes in the elderly that occur as a result of disease or aging. Any of these factors singly or in combination can adversely affect postural control. Therefore, balance is an important issue in the elderly and must be measured well to provide a good assessment, to plan appropriate interventions and to evaluate the outcome.

2.4. Assessment of balance

2.4.1. Laboratory measures of balance

Postural sway has been measured by using simple swaymeters attached to the shoulders (Sheldon 1963) or waist (Overstall et al. 1977), as well as by sophisticated laboratory equipment that measure excursions of the centre of pressure with force plates. The centre of pressure, that is the centre of the distribution of the total force applied to the supporting surface, moves according to the excursion of the centre of gravity of the body and the distribution of muscle forces required to control or produce the movement. Measurements can be made of the amplitude of sway or speed of the centre of pressure excursions; or they may be expressed as force measurements which include an assessment of the acceleration of the centre of gravity.

The accuracy of the centre of pressure displacements is considered to be high. Few studies have examined the stability of subjects' performance. Holliday and Fernie (1979) measured the spontaneous postural sway of 29 normal healthy adult subjects, aged 22 to 64, on 15 consecutive days. The value of the speed of sway for each day represented the average of four 1-minute recording periods. Despite the use of this daily average, there was a marked learning effect from repeated testing corresponding to a 31% reduction in the mean speed of sway during 15 sessions.

The precision of force plates in measuring postural sway is relatively easy to establish. It is more difficult to examine the validity of the test as a measure of balance. Some evidence of validity is provided by studies showing a moderate association between postural sway and clinical tests of balance (Berg et al. 1992; Lichtenstein et al. 1990). A stronger correlation is not expected because the clinical tests primarily address postural adjustments to voluntary movements whereas the laboratory measures examine the ability to maintain a position or adjust to external disturbances.

Further evidence of validity is provided by a study in which a swaymeter was able to detect changes in amplitude of postural sway in accordance with the expected pattern, duration and dosage of the administration of a sedative in healthy young and elderly volunteers (Swift 1984). However, in a separate comparison, there was no correlation between drug plasma concentrations and postural sway in habitual (1 month to 15 years) sedative users (Swift 1984). The absence of a relationship may reflect the ability of individuals to adapt to longterm use or that amplitude of postural sway is too insensitive a measure to detect subtle chronic changes whereas they could detect acute responses to the medications.

In examining possible reasons for the inability of postural sway to differentiate between treatment and control groups in studies evaluating the effects of exercise (Crilly et al. 1989; Lichtenstein et al. 1989), it is worthwhile to examine the underlying assumptions of measures of postural sway.

The first assumption is that a greater amplitude of sway is indicative of less ability to maintain balance. This hypothesis is not necessarily correct. Gauthier-Gagnon and colleagues (1986) found that amputees recently fitted with a prosthesis swayed less than normal subjects. As suggested by the authors, subjects may have compensated for the change in proprioceptive input by fixating visually or by stiffening their posture while concentrating on the task of standing. The normal subjects may have allowed themselves to relax and sway over a greater amplitude because they found the task to be very easy.

In an effort to better discriminate between subjects of varying ability, researchers have recorded sway of subjects while in positions of increasing difficulty: stepping with one foot ahead of the other; standing with one foot directly in front of the other (tandem); and standing on one leg (Goldie et al. 1989; Lichtenstein et al. 1989). This solution may not be practical for evaluating geriatric subjects as the ability to hold the positions, in particular one legged stance, has been shown to decline with age (Era and Heikkonnen 1985; Potvin et al. 1980).

The second basic assumption of postural sway measures is that subjects with a faster speed of sway have greater balancing difficulty. Hence, when two individuals have the same amplitude of sway, the one with the faster speed is working harder to maintain his or her posture. In contrast to the study which compared amputees and normal subjects on the basis of amplitude of sway (Gauthier-Gagnon et al. 1986), Fernie and Holliday (1978) were able to demonstrate that amputees had more difficulty with balance on the basis of a faster speed of sway.

In a cross sectional study of 31 elderly subjects, Berg and associates (1992) found that speed of sway during quiet standing showed the highest correlation with the Balance Scale and was the only laboratory measure to discriminate between groups based on their use of a walking aid. A multiple regression analysis showed that inclusion of both amplitude and speed of sway in response to the pseudorandom movements of the platform explained more of the variance in the Balance Scale scores than did either parameter alone. Moreover, there was little relationship within each subject between the amplitude and speed of sway in response to the perturbations.

Force plates are also used to map the base of stability by asking the subject to lean as far as possible forward or backward (Dettmann et al. 1987; Murray et al. 1975). A larger area indicates greater balancing ability because it brings the individual closer to the edges of the base of support and requires greater effort to prevent a loss of balance. A larger area also represents a subject who can control his or her centre of gravity within a greater diameter. This method of assessment incorporates the subject's self perception of how far it is possible to lean safely. As could be expected, the area of stability of stroke patients is related to their motor performance and functional status, and is smaller for hemiplegic patients when compared to normal subjects (Dettmann et al. 1987).

Duncan and associates (1990) assessed reliability of voluntary centre of pressure excursions in a forward direction. Measures were taken when subjects were asked to reach forward with their arms. The test-retest reproducibility of the performances of 14 elderly subjects tested one week apart was .52 as estimated by the intra class correlation coefficient (ICC).

This type of test has lead to commercial equipment which provide visual feedback to subjects wanting to learn how to control their centre of gravity, to reduce sway, or to assume a more symmetrical stance.

The disadvantages of the laboratory measures are the expense of the equipment, the need for well trained personnel, and the questionable functional relevance of the tests. Standing for a prolonged period and travelling to a laboratory may also represent an unnecessary burden on an elderly subject.

2.4.2. Clinical measures of balance

The evaluation of balance in the clinical setting is generally subjective. To grade a patient's balancing ability a professional may give a global rating of poor, fair or good. Progress may be noted by stating the level of assistance or supervision required for a specific activity. Several clinical scales are available and used to varying degrees in clinical practice and research.

2.4.2.1. Fugl-Meyer Balance Sub-Scale

Descriptive methods to denote the presence or absence of equilibrium reactions and the quality of postural symmetry have been advocated for patients with neurological disorders (Bobath 1965; 1970). This approach was modified by Brunnstrom (1970), expanded by Fugl-Meyer et al. (1975) and incorporated as part of a quantitative assessment of physical performance for patients with cerebrovascular accidents (CVA). The validity of the instrument as a whole has been tested only on the stroke population (Duncan et al. 1983; Fugl-Meyer 1975). Moreover the balance section has not been tested or used independently. This section of the Fugl-Meyer instrument grades seven tasks, three in sitting and four in standing. A three point ordinal scale ranging from 0, cannot be performed, to 2, can be fully performed, is applied to each task. The subject is timed while standing on two feet, while on each foot alone, and while sitting unsupported. In addition, the examiner pushes the subject from side to side to observe the postural reactions in the sitting position. Clinically it seems unlikely that this scale is sufficiently sensitive for use as a measure of progress or deterioration. In particular, there may be a large delay between mastery of the easier items and the ability to stand on one leg. During this time, actual changes in status are likely left undetected.

2.4.2.2. Balance Coding

A more comprehensive indicator of balance was developed by Gabell and Simons (1982). They used a code to designate a geriatric patient's ability to withstand static, sagittal and rotational stresses and thus show his or her level of competence. Static stress has six levels ranging from 0, for unsafe in sitting, to the maximum of 6, when the subject can stand steadily 20 seconds with eyes closed without any mechanical aid and with one foot directly in front of the other. The levels follow the principles of body mechanics by asking the subjects to gradually diminish the base of support. Rotational stress is applied by asking the patient to turn his head, and to turn his body 360 degrees.

Each of these tasks is graded independently, A and B. Finally an X is awarded to the person who can safely stand up from a sitting position and stand still for twenty seconds. Rather than give a cumulative score, an aggregate profile is given which categorizes the patient.

There has been little documentation of the measurement properties of the scale; however, it is among the first to concentrate on problems specific to the elderly. Moreover, the mean scores did show a gradient from low to high for geriatric in-patients, geriatric out-patients, and normal healthy elderly. This finding may be interpreted as evidence for the "known groups" type of criterion validity.

2.4.2.3. Tinetti Balance Sub-scale

Tinetti and associates (1986) developed a Fall Risk Index to identify individuals with a propensity for falling. Within this Fall Risk Index, the single best predictor was a mobility score incorporating measures of gait and balance (Tinetti 1986). The balance section contains thirteen manoeuvres of which some are graded dichotomously, can/cannot perform. The other items are scored 0, 1 or 2 to denote quality of performance. The total score can range from 0 to 24 points.

Some data for reliability of the mobility score are available. Inter-rater reliability, as scored by a physician and nurse simultaneously, showed agreement on 85% of the items and a total score that differed by less than 10%. Information for test-retest reliability has yet to be documented.

Evidence for concurrent criterion validity is the association demonstrated between the Balance subsection of the mobility scale and laboratory measures of postural sway (Lichtenstein et al. 1990). A prospective study on falls in the elderly showed that this mobility score was the best single predictor of recurrent falling (Tinetti et al. 1986). Similarly selected items from the mobility scale were included in the final mathematical model predicting falls in a community based study (Tinetti et al. 1988). Given the emphasis on predicting the occurrence of falls, the precision of the instrument in discriminating differences between individuals or within the same person at two points in time has not yet been assessed.

2.4.2.5. (Timed) Get Up and Go

Mathias and associates (1986) have devised a "Get Up and Go" test, also based on functional performance. Subjects are asked to rise from a chair, stand still momentarily, walk toward a wall, turn and then return to the chair. They receive a score based on the quality of their movements. It is a practical and simple test but by no means a comprehensive measure of balance. Individuals are given only a subjective global rating ranging from 1 (normal) to 5 (abnormal). Because of its global scoring system, this test can likely only be used to screen patients to determine if they need closer attention. It would not be suitable to monitor patients over time or to detect small changes in their status.

Podsiadlo and Richardson (1991) have used the same manoeuvres but record only the time to complete the test. The total scores have shown a strong association with a subjects' functional status as measured by the Barthel Index of Daily Living (1965) and balancing ability (Berg et al. 1989). The reliability was excellent for inter and intra-rater agreement (ICC .98 for each).

The change in the scoring method also caused a rethinking of the underlying concept. Currently the test is considered a test of basic functional mobility rather than a measure of balance (Podsiadlo and Richardson 1991).

2.4.2.6. Postural Stress Test and Maximal Load Test

A quantitative method suggested for assessing the ability to respond to external forces is the "postural stress test" (Wolfson et al. 1986). This test involves using a pulley weight system to deliver a backward pull on the hips while the subject is standing. The videotaped performance of the patient's response to three different weights is scored based on the sequence and type of response to each weight. The test was able to differentiate among three groups known to differ: young subjects, healthy elderly and elderly subjects with a history of falls. Inter rater reliability was reported to be .99, but as this estimate was based on Cronbach's alpha, it is difficult to interpret the finding. When the test was repeated on consecutive days, the ICC was .83 showing moderately high reproducibility (Hill et al. 1990).

The "maximal load" test (Lee et al. 1988) requires subjects to maintain a standing position against static loads applied at the waist from behind. Maximal loads are recorded as a percentage of body weight at the point where the subject can no longer hold the initial standing position.

Limitations of the "maximal load" and "postural stress" tests are the need for equipment, the lack of portability and the questionable appropriateness of such destabilizing forces. They do not replicate a functional situation. Given the evidence that balance responses may be specific to the methods of perturbations (Duncan et al. 1990) and the low correlation of the "maximal load" test with clinical assessment (Lee et al. 1988), it may be advisable to test in more functional ways.

2.4.2.7. Clinical Test for Sensory Interaction and Balance (CTSIT)

A new approach to evaluating balance has been suggested by Shumway-Cook and associates (1986;1991; Horak 1987). They have incorporated the laboratory methods of Nashner and associates (1982;1987) into a systematic clinical evaluation. Subjects are asked to stand for 30 seconds on a firm surface with eyes open, blindfolded, and wearing a head dome to give inaccurate visual information. This sequence is repeated with the subject standing on a piece of foam to give inaccurate somatosensory inputs. Any sway is assessed on a subjective rating scale. In addition, the patients' movement strategies, in response to external forces and during voluntary arm raises, are observed.

The observations and data from this test bring clinical practice parallel to laboratory research. Clinicians are encouraged to test patients for the use of ankle, hip or stepping strategies. While it is not yet known what constitutes a normal range of effective strategies for every situation, it is helpful to determine if subjects have a range of strategies at their disposal and what factors may account for the absence of an effective strategy. For example, the ankle strategy will be ineffective if there is a diminished passive range of motion at the ankle.

A limitation of the CTSIT is that there is no global score that describes the subject's ability. It is a comprehensive assessment in that it seeks the origin of the problem and assesses the components of various systems. To date, it has not been used as an outcome measure. Further work is needed before clinicians can determine when normal variability becomes an abnormal strategy and functionally ineffective.

2.4.2.8. Functional reach

Duncan and colleagues (1990) have suggested functional reach as a new clinical measure of balance. This single test refers to the distance a subject can reach his or her arm forward while standing. It is conceptually similar to laboratory tests of the excursion of the centre of pressure while subjects voluntarily lean forward. The same task is included in the Balance Scale (Berg et al. 1989) except it is scored differently and the distance reached is measured with a normal ruler. Functional reach is assessed as the average of three measurements using a level yardstick secured to the wall at shoulder height. The coefficients for test retest and inter-rater reliability of the clinical yardstick measurements were .92 (ICC 1,3) and .98 (ICC 1,3), respectively.

The validity of the test was assessed through comparisons with the electronic yardstick and the centre of pressure excursions while subjects leaned forward (Duncan et al. 1990). Both variables showed a moderately high association with the clinical yardstick measure. The distance reached was also related to age and anthropometric body measurements.

This test is simple, easily transportable, and has demonstrated good measurement properties to date. Given that it is a single test, it may not be sufficiently comprehensive to be used as a descriptive measure. In addition, the relationship between functional reach and more global performance has not been carefully explored.

Summary

There is little consistency in the use of measures of balance in clinical practice or research. Laboratory measures are expensive, require specialized equipment, and do not directly relate to functional ability. Clinical measures have become more quantitative, but each has limiting features and unanswered questions concerning its measurement properties.

2.5. Chapter Summary

Balance is an integral component of motor control and movement. There are many reasons for its impairment in the elderly which predisposes this population to fall or limits their functional performance. We need a measure of balance with good measurement properties to describe the extent of dysfunction in various sub-groups, assist in setting priorities for prevention, develop intervention strategies and assess their effectiveness. A performance-based measure of balance that meets multiple requirements of clinical practice and research would promote a continuity of information and contribute to developing a greater depth of knowledge in the area.

The existing scales have not as yet demonstrated the necessary requirements. Relative to the other measures, the Balance Scale shows greater promise of serving multiple purposes. It is portable, easy to administer, requires only a ruler and stopwatch as equipment and takes only 10-15 minutes to complete. In fact, administering the scale does not take additional time because clinicians include the items in their routine evaluations. The Balance Scale gives them the opportunity to quantify their observations. Moreover, the Balance Scale is the only measure of balance to have undergone formal content development and testing of its internal consistency. Although not yet tested, the independence- dependence continuum of the response choices for each item should permit a fairly precise discrimination between subjects of varying ability and within subjects changing over time. These qualities are important for an outcome measure.

Before advocating its use, the Balance Scale requires testing of its reliability in realistic clinical situations and further evidence of its validity. The specific objectives and methods used to assess the performance of the Balance Scale are presented in the next chapter.

Chapter 3 THE MEASUREMENT STUDY

3.1. Objectives

The general objective of this thesis is to examine the measurement properties of the Balance Scale, a measure which has demonstrated excellent reliability and content validity in preliminary testing, but which requires more extensive testing to assess its performance. Several strategies to examine the performance of this scale were developed and tested in two groups of subjects: residents of a home for the elderly and patients admitted to general hospital with acute stroke. The present measurement study has three components: Validity Study I, Validity Study II and the Reliability Study. Validity Study I examines the concurrent and predictive criterion validity of the Balance Scale in the group of elderly residents. Validity Study II examines the construct validity and ability of the Balance Scale to monitor changes in the status of patients following the onset of stroke. The third component assesses the reliability using both elderly residents and acute stroke patients. The specific objectives are grouped to correspond with the three inter-related studies.

3.1.1. Specific objectives of Validity Study I

1. To determine if Balance Scale scores can differentiate among groups of elderly individuals on the basis of the type of mobility aid used ("known groups" technique of concurrent criterion validity).
2. To determine the degree of association between Balance Scale scores and the global ratings of good, fair and poor balance given by clinicians (concurrent criterion validity).
3. To assess the association between Balance Scale scores and the occurrence of falls in the following year (predictive criterion validity).

3.1.2. Specific objectives of Validity Study II

1. To assess the degree of association between Balance Scale scores and motor performance and functional status scores at various points in the recovery of stroke patients (construct validity).

2. To determine if Balance Scale scores are able to detect changes in the status of stroke patients to the same degree as the Barthel Index (responsiveness).
3. To determine whether Balance Scale scores are able to discriminate between groups of stroke patients defined by where they are residing at the time of the evaluation ("known groups" type of concurrent criterion validity).

3.1.3. Specific objectives of the Reliability Study

1. To assess the inter-rater reliability between pairs of observers independently rating the same subject.
2. To assess the intra-rater reliability of the same observer rating the same subject at two points in time.
3. To assess the internal consistency of the Balance Scale when used with elderly residents and the acute stroke patients.

3.2. Methods

The strategies used to examine the performance of the Balance Scale are explained separately for each of the three sub-studies. Validity Study I describes the methods used in the home for the elderly to examine the criterion validity of the Scale. Validity Study II, outlines the approaches employed to assess the criterion and construct validity and responsiveness of the instrument in stroke patients. The Reliability Study presents the methods of assessing reliability. Although it is a prerequisite to validity, reliability is addressed last because subjects in this section represent sub-groups of the elderly and the stroke patients in the two validation studies.

3.2.1. Validity Study I

The three specific objectives of Validity Study I address criterion validity. The first two, as examples of concurrent criterion validity, examine the association between Balance Scale scores and walking aids and clinicians' global judgments of balance (good, fair, poor) at a given point in time. However, a longitudinal design was required to assess predictive criterion validity as determined by the relationship of Balance Scale scores with future falls in the elderly population.

3.2.1.1. Locus of the study

The Griffith-McConnell Residence in Montreal is a home for 350 senior citizens. It is divided into sections based on level of independence and degree of required assistance or supervision. While there is an overlap in ability, residents of the McConnell are considered to be the most active and independent. Their outside activities vary from short outings accompanied by family or friends to unrestricted use of public transportation in the city and travel outside the province. Residents in the Griffith tend to be independent indoors but require accompaniment when outdoors. Residents in both the McConnell and Griffith sections attend their respective dining rooms for meals. The Annex offers greater nursing supervision and assistance with bathing. Residents in this area are independently mobile but do not attend the main dining rooms. In addition, the residence has an infirmary for residents requiring nursing care and a recently established special care unit for individuals with cognitive impairment.

There are no physical or occupational therapists employed at the Residence, but the Women's Auxiliary and the recreational therapist plan a variety of diversional activities. The residents are active in their own affairs with a Resident's Council that has representation from all sections. Collaboration for this study was obtained both from the Ethics Committee of the Board and the Residents' Council (Appendix 3).

3.2.1.2. Subjects and methods

A longitudinal study was conducted, wherein residents were recruite[illegible] the Griffith, the McConnell and the Annex and followed for one year. The eligibility criteria were: age 60 years and older, medically stable, independently mobile with or without a walking aid, and willing to participate in the study. Medically stable referred to the absence of a known medical problem that would interfere in the subject's ability to complete the year of follow-up.

A formal sample size calculation was not made. However, based on incident reports of falls in the residence for the years 1986-87 and statistics stating that one-third to one-half of all persons age 65 and older fall at least once a year (Kellogg International Work Group 1987; Nickens 1985), it was estimated that 30 to 50 individuals out of 100 would fall in the year of follow-up. Therefore, a sample size of 100 would be sufficient to permit a comparison between fallers and non-fallers. In addition, because of a mortality rate of 16% in the previous year at the Residence, it was deemed necessary to recruit 115 persons.

Efforts at recruitment began at the Resident's Council Meeting and with posters strategically located around the residence. Two part-time research assistants, both registered nurses, scheduled their time so that one or both were on site at least four days of the week. Their continuing presence and the central location of the office space allotted to the study permitted close contact with staff and residents, who then encouraged others to join the study.

One research assistant met with each potential participant to explain the purpose and procedures of the study. After obtaining informed consent, she recorded baseline sociodemogaphic and clinical information from the resident and the chart, and asked about current medications, presence of medical conditions, and history of falls in the previous three months. The documentation pertaining to screening procedures, informed consent and patient profile forms are included in Appendix 3.

She administered the Mini-Mental State Exam (Folstein et al.1975), a screening tool which has been used in geriatric populations. A copy of the test is included in Appendix 2. The 11 questions on the Mini Mental State Exam (MMSE) address orientation, memory, language, calculation, attention and spatial ability. Total scores range from 0-30. A score below 18 indicates definite impairment whereas a score between 23 and 18 is suggestive of mild cognitive dysfunction (Bleecker et al. 1988; Tombaugh and McIntyre 1992). All subjects signed the consent form; however, an additional form explaining the subject's participation in the study was sent to a family member if the subject scored less than 24 on the MMSE.

A research assistant also administered the Barthel Index of Daily Living (Mahoney and Barthel 1965). The Index measures functional levels of independence by assessing 15 items related to self-care and mobility (Granger et al. 1976; 1977; 1979; 1979). Each item is scored by determining whether the patient can perform the requested activity independently, with assistance or supervision, or not at all. The scores for each item are summed and the total can range from zero (complete dependency) to 100 (independence in terms of personal care). The Index can also be considered in terms of its two sub-scales, Self-Care (0-53) and Mobility (0-47).

A copy of the Barthel Index is included in Appendix 2. In addition, the Index is discussed in greater detail in the instrumentation section of Validity Study II, in which it serves as the criterion measure. In Validity Study I, both the Mini-Mental State Exam and the Barthel Index were primarily used as descriptive measures.

Balance Scale evaluations were performed by independent evaluators at baseline, three, six and nine months following entry to the study. The independent evaluators were either occupational or physical therapists. They were given a brief training session in which they reviewed the items and had an opportunity to practice on each other. They came to the Residence for specific appointments, scheduled by the research assistants.

In addition, at entry to the study all subjects were asked about the type of walking assistance they generally used. The four levels of ability were defined as independent without any aid, uses a cane when outdoors only, uses a cane and lastly, requires a quad cane, walker or wheelchair to get around.

To address the second type of concurrent criterion validity, the research assistant asked a staff member familiar with the patient, but blind to the results of the other measures, to rate the subject's balancing ability as good, fair or poor. This rating was obtained within 24 hours of the independent evaluator's assessment of balance. This procedure was repeated at three, six and nine months to assess the consistency of the relationship. At each time, the residents were also asked to judge their ability to balance as good, fair or poor.

Falls were ascertained through verbal reports from the staff and other residents, official incident reports, the medical charts, and self reports from each subject. Participants were asked to inform the research assistants when they had a fall. Details of the fall were obtained during a private interview with the resident. Reminders to report a fall were given at Resident's Council meetings periodically throughout the two-year study period. Additionally, each participant was asked about falls at each three-month evaluation and upon completion of the year's follow-up. Finally, at the end of the study, all charts and incident reports were reviewed.

3.2.1.3. Analysis

The accuracy of the coded information was assessed in several ways. The patient profile sheets were compared to the original documents in a sample of 20 charts selected using a table of random numbers. Once entered into the computer, all files were checked against the profile sheets. Finally, descriptive statistics were used to verify that the range of scores were consistent with the expected values.

Descriptive statistics also provided information on the socioeconomic and baseline clinical characteristics of the subjects before proceeding to address the three specific objectives.

The first objective was to determine if Balance Scale scores could discriminate among groups known to differ in the use of walking aids at entry to the study. A frequency distribution of the use of aids and the mean Balance Scale scores for each level of aid were examined. A one-way analysis of variance was used to compare the mean Balance Scale scores for each group. In addition, an analysis of variance using polynomial contrasts permitted an assessment of a linear relationship between the Balance Scale and the levels of aids, coded: no aids (1), cane outdoors only (2), cane (3) and walker (4).

Other factors potentially influencing the use of mobility aids were also examined. They included mental status, Barthel mobility sub-scale, age, orthopaedic, neurological and rheumatological diagnoses. An analysis of variance using these covariates was employed to assess whether the difference in mean Balance Scale scores remained statistically significant.

The second objective was to determine the degree to which Balance Scale scores were related to the clinicians' concurrent global ratings of balancing ability. A frequency distribution of the global ratings (good, fair, poor) and the mean Balance Scale scores for each global rating category were examined at each evaluation point. A one-way analysis of variance was performed to assess whether the observed differences in means between the three categories were greater than expected by chance alone. Polynomial contrasts were used to assess the presence of a linear trend in the means for each category, coded as: good (3), fair (2) and poor (1). In addition, given that the global ratings may reflect the functional status or age of the subject rather than balance, the mean differences in Balance Scale scores between the groups were examined using Barthel Index scores and age as covariates.

The strength of the relationship between the Balance Scale scores and the global judgments was tested using Spearman Rank Order Correlation Coefficent (rho). The consistency of the association was verified by repeating the analysis for each pair of ratings (3, 6 and 9 months). Correlation coefficients were also used to assess the relationship between Balance Scale and the self report of balancing ability at each point in time. The levels of self ratings were good, fair, poor.

The third objective was to determine the relationship between Balance Scale scores and the occurrence of falls in the year of follow-up. The clinical and functional characteristics of the subjects at entry to the study were examined relative to their fall status at the end of the study: non-faller, single and multiple time faller. This comparison permitted an inspection of variables potentially influencing the relationship between balance and falling.

The stability of the Balance Scale and Barthel scores for elderly residents during the study was assessed by a repeated measures analysis of variance examining within subject differences over the four evaluations.

The relative risk of falling at any point in the year was determined by using a score of below 45 on the Balance Scale at the baseline evaluation as the risk factor. This cut-off was based on the experiences of clinicians familiar with the Scale who believed that patients scoring below 45 required greater assistance or supervision than those scoring 45 and above.

This analysis was repeated comparing the risk of falling two or more times compared to once or not at all. It was expected that the relationship between multiple falls and balancing ability would be stronger than for a single fall. Previous work (Berg et a'. 1992) had indicated that single fallers and multiple fallers were different in terms of baseline clinical and functional characteristics.

The strength of the association between the initial Balance Scale scores and the fall status at the end of the year was also assessed by the Spearman's Rank Order Correlation Coefficient. Fall status was ranked as 0, 1 or 2 or more falls within the year.

Logistic regression was used to assess whether the Balance Scale scores can be useful in combination with other variables in predicting falls. Variables were considered for inclusion in the model based on factors associated with falls in previous studies or if indicated by the comparison of the baseline clinical and functional variables between the fall status groups. Potential variables were age, history of falls in the past three months, cognitive impairment, number of diagnoses, current use of medications and visual impairment. The equations were formulated using at least 1 fall/0 falls, alternately with 2 or more falls/0 or 1 fall, as the dichotomous outcome variables. Variables were entered into the model based on the clinical rationale, avoiding any combination of variables that were strongly intercorrelated. Logistic regression using stepwise backward procedures and the log likelihood ratio test were used to determine which potential variables made a useful contribution to the model.

Additionally, mean Balance Scale scores at the beginning of each interval were examined among non-fallers, single fallers and multiple fallers. It was expected that the number of individuals falling in every three month period would be relatively small and not amenable to logistic regression. Moreover, considering each interval separately would not be an independent analysis. Nonetheless, there should be a trend of higher scores in the non-fallers and lower scores in the recurrent fallers that remains relatively consistent during the intervals. A one-way analysis of variance was used to compare the mean Balance Scale scores of the non-fallers, single and multiple fallers.

3.2.2. Validity Study II

The objective of this part of the measurement study was to assess the validity of the Balance Scale in a population of acute stroke patients. Three different strategies were used to assess how well the scale performed in clinical situations. The relationship between balance and functional status and motor performance was explored as a method of construct validation. The responsiveness of the scale was examined using the Barthel Index as the external criterion denoting a change in clinical status. Concurrent criterion validity was assessed by comparing mean balance scores of groups determined by the location of the follow-up evaluation.

3.2.2.1. Locus of the study

The Royal Victoria Hospital (RVH) and the Montreal General Hospital (MGH) are large acute care institutions situated in central Montreal. They are both affiliated with McGill University, active in teaching and research and offer most medical and surgical specialities. The RVH is closely associated with the Montreal Neurological Institute (MNI), a combined care and research centre specializing in neurological conditions. Stroke patients are admitted to these institutions for acute care, after which they may be referred to a convalescent hospital for rehabilitation, or in the case of the RVH transferred to the Geriatric ward. Each institution receives services from a Department of Physical Therapy. The Department at the RVH serves both the RVH and the MNI. The MGH has its own Department of Physical Therapy.

Secondary institutions to which stroke patients are referred for rehabilitative services include the Julius Richardson Convalescent Hospital, the Catherine Booth Hospital Centre, the Montreal Convalescent Hospital Centre and the Jewish Rehabilitation Hospital. Collaborative agreements were made with each institution whereby a specific individual, designated as the liaison for the study by the Director of the institution, was approached at the time of a follow-up assessment. Appropriate arrangements were made with that individual and the patient to schedule and complete the follow-up visit.

In addition to these institutions, patients were followed at Centre Hospitalier Côte des Neiges, L'Hôpital Marie Clarac, as well as in their homes. In these situations, specific agreement was negotiated for each individual patient.

3.2.2.2. Subjects and methods

A longitudinal study was conducted, wherein stroke patients were recruited from the general hospitals and followed for 12 weeks. The eligibility criteria were: age 40 years and older, medically stable, admitted to hospital with an acute stroke of less than 14 days duration, showed evidence of motor impairment, and lived in the greater Montreal area. Exclusion criteria included medical factors that could interfere with rehabilitation such as blindness, lower extremity amputation, advanced Alzheimers disease and history of functional dependency prior to the onset of stroke. An estimate of the subject's previous level of functioning was obtained from the admission notes or from family members. Based on this information, a pre-admission Barthel Index score was estimated, and subjects with scores below 40 were excluded from the study. Documentation concerning the screening process, consent forms, and patient profile sheets is included in Appendix 3.

Three physical therapists acted as part-time research assistants within their own institution. They were responsible for determining the eligibility of subjects referred by other rehabilitation staff, assuring the consent of attending medical staff, explaining the study to the potential subjects, and obtaining informed consent. If there were possible comprehension or speech problems, the patient's family was asked to sign a form stating that they were aware of the patient's participation in the study.

The research assistants collected baseline sociodemographic and clinical data from the patients' charts or through a brief interview. They were also trained to administer the Barthel Index and the Fugl-Meyer Motor Performance Scale (Fugl-Meyer et al. 1975). They performed these measures within 24 hours of the Balance Scale assessments made by the same independent evaluators, who participated in Validity Study I. The baseline evaluations were completed 2 weeks post onset of stroke. At 4, 6, and 12 weeks post stroke. More frequent evaluations were made in the first six weeks because this is the period in which stroke patients should show the greatest change. The research assistants retained responsibility for all scheduling of appointments throughout the study period; however, the Project Director completed the repeat assessments of the Barthel Index and the Fugl-Meyer Motor Performance Scale once the patient was discharged from the general hospital.

3.2.2.3. Instrumentation

Functional status

The Barthel Index has been described briefly in the previous section (Validity Study I). Although it was used to assess the residents of the home for the elderly, it is more widely recognized as a measure of functional status in stroke patients and has been used internationally in a wide variety of studies of both acute and chronic stroke patients (Chino 1990; Granger et al. 1988; Hewer 1990; Reding 1990; Wood-Dauphinee et al. 1990).

Granger and associates (1979) reported that 60 represents a cut-off between dependence and assisted independence, 40 signifies dependence, and 20 total dependence. The self-care section is scored 0-53; 20 of these points are awarded if the subject is continent. The other 33 points pertain to eating, drinking, dressing and grooming. The mobility section is scored from 0-47. The items include toiletting, transferring, getting in and out of a bathtub or shower, stairs and walking. Although essential to independent living, the items address ten very basic activities and thus do not discriminate well in high level subjects.

Barthel scores have predictive validity in terms of survival (Wylie et al. 1964;1967), functional recovery (Granger et al.1975;1977;1979; Wylie 1967; Hertanu et al.1984) and discharge (Granger et al.1977). Barthel scores correlate highly with other daily living indices (Donaldson et al.1973; Gresham et al.1980) and medical assessment of clinical improvement (Wylie 1967). Dettmann and colleagues (1987) determined that scores on the Barthel Index are related to selected laboratory measures of postural sway and stability. It has demonstrated a test-retest reliability of 0.89 and an interrater reliability of 0.95 (Gresham et al.1980).

The Barthel Index is moderately sensitive to changes in a patient's status (Donaldson et al.1973). As further evidence of responsiveness, the Barthel Index demonstrated a larger effect size for detecting a treatment effect in acute stroke patients than measures of neurologic status, stroke severity and motor performance (Wood-

Dauphinee et al. 1990). It may, however, have limited responsiveness for discriminating changes in high functioning subjects because it addresses only basic activities of daily living. Given the wide range of ability expected in the present study, this limitation is not of concern.

Motor Performance

Fugl-Meyer and associates (1975) have developed an assessment tool which quantifies recovery of motor performance post stroke. The instrument was constructed under the assumption that recovery of motor function follows a stereotyped sequence of motor events. A 3-point scale is applied to each item with a maximum score for all components of 226. Sub-scale scores for the lower extremity, the upper extremity, balance ability, sensation and range of motion may also be calculated.

The balance section grades seven tasks, three in sitting and four in standing. Subjects are assessed for their ability to sit unsupported and react appropriately when the examiner pushes them side to side while sitting. In the standing position subjects must be able to stand supported, unsupported and on each leg alone. Although, they contain three similar tasks, the Balance Scale and the Fugl-Meyer sub-scale score each differently.

The overall concurrent validity (Fugl-Meyer 1976; Kusoffsky et al.1982, Wood-Dauphinee et al.1990), the predictive validity of the lower extremity staging system (Clarke et al.1983) as well as intra-rater and inter-rater reliability (Duncan et al.1983) of the scoring system have been established to be within acceptable limits for use in a clinical study. In addition, the instrument is responsive to fairly small changes in patient status (Fugl-Meyer 1976). Badke and Duncan (1983) also determined that the lower extremity subscale reflected proportional electromyographic activity when hemiplegic subjects were asked to balance on a moving platform. The instrument is easy to administer by trained personnel and requires no complicated equipment. A copy of the Balance, Arm and Leg sub-scales is included in Appendix 2.

3.2.2.4. Analysis

To assess the accuracy of the information entered in the computer files, each file was checked against the patient profile sheets by two persons. In addition, descriptive statistics were examined to verify that the range of scores was consistent with the expected values.

Descriptive information on the sociodemographic and baseline clinical characteristics of the sample was examined before proceeding to the three specific objectives in Validity Study II. The first objective was to assess the degree of association between the Balance Scale and the motor performance and functional status scores at various points in the recovery of stroke patients. Pearson's Correlation coefficient was used to assess the strength of the linear relationship between the Balance Scale and the Barthel Index and the Fugl-Meyer Scale at entry to the study, 4, 6 and 12 weeks post onset of stroke. Additionally, the pattern of change in the Balance Scale, the Barthel Index and the Fugl-Meyer Motor Performance Scale were examined graphically at each evaluation point.

The covariation of the three measures also relates to the second objective concerning the responsiveness of the Balance Scale to clinical change in status. Similar profiles of change among the three measures would provide indications of the responsiveness of the instrument, as the Fugl-Meyer Scale and to a greater degree the Barthel Index have both been shown to be sensitive to clinical changes (Fugl-Meyer et al. 1975; Wood-Dauphinee et al. 1990; Wylie 1969).

A three step procedure was used to compare the responsiveness of the Balance Scale relative to the Barthel Index. First, an analysis of variance examining within subject variation over time was used to verify that the Barthel Index was able to detect a change in the status of the subjects over the 12-week follow-up period. Polynomial contrasts were used to assess whether the trend in time was linear, quadratic or cubic. Second, this analysis was repeated in the same patients using the Balance Scale as the dependent variable.

Once it was demonstrated that both the Barthel Index and the Balance Scale were able to detect an alteration in status over the 12-week period, the third step was to determine if the changes in the Balance Scale were of the same magnitude as that of the Barthel Index. To address this question, a repeated measures analysis of variance examining within subject changes over time was performed using the difference between the Barthel and the Balance Scale scores for each subject at each evaluation point as the dependent variable. For this analysis the Balance Scale scores were converted to a score out of 100. The absence of any effect of time on the difference score would indicate that the changes in the two measures were of the same magnitude, and that the Balance Scale was at least as sensitive as the Barthel Index to detecting change in the status of stroke patients.

This same three step procedure was repeated to examine the responsiveness of the Fugl-Meyer Arm and Leg and Balance Sub-scales relative to the external criterion, the Barthel Index.

In addition, these analyses were repeated in three clinical sub-groups of the stroke population based on their Barthel Index scores at entry and at the end of the study. The sub-groups were defined prior to the analyses. The low and high functioning groups were defined as individuals with Barthel Index scores below 40 and above 60, respectively, both at entry to the study and at the end of 12 weeks. The middle group included the rest of the 60 patients in the study.

The response to change within these clinically defined sub-groups was examined in the same manner as for the whole population. It was expected that the middle group would demonstrate the greatest rate of change but that within each sub-group at each point the c.variation between the Balance Scale, Barthel Index and the Fugl-Meyer Scale would remain high.

The third specified objective was to compare mean Balance Scale scores between groups that had been formed according to the location of the follow-up evaluation. While other factors may influence discharge home or to rehabilitation facilities, it is assumed that there would be differences in ability between individuals living at home, at a rehabilitation hospital, and in the general hospital. A one-way analysis of variance was performed to test whether the differences in mean scores was greater than expected by chance at each of the follow-up evaluations.

3.2.3. Reliability Study

As a prerequisite to validity, reliability was assessed in the preliminary study of the Balance Scale (Berg et al. 1989). In the present measurement study, it was examined using subjects from both validity studies. Consequently, to avoid repetition of the methods, the reliability study is presented last.

3.2.3.1. Inter and intra rater reliability

Subjects and methods

The senior matrons and nurses at the Griffith-McConnell Residence were invited to participate in the reliability study. To examine inter-rater reliability, they were asked to administer the Balance Scale to residents with whom they were familiar, within one week of the corresponding assessment of the independent evaluators. To assess intra-rater reliability, the Balance Scale was administered twice by the same person, at least one week apart. The residents were asked to perform these additional tests toward the end of the study period so as not to compromise their participation in the scheduled evaluations. Additionally, we did not wish the knowledge of the Balance Scale to influence the global judgments of balance made by the staff.

In the stroke study, nurses, and occupational and physical therapists were asked to administer the Balance Scale within 24 hours of the independent evaluator. The evaluations were carried out in both the general hospital and in the rehabilitation facilities.

Analysis

To assess the accuracy of the computer files, the information entered into the computer was checked against the patient profile forms. Additionally, descriptive statistics were examined to verify that the range of scores were consistent with expected values.

Baseline clinical and sociodemographic characteristics were compared between the sub-group of the elderly residents and stroke patients in the reliability study and those in Validity Studies I and II. In addition, the mean and range of values of each pair of ratings was examined within each sub-group.

Inter observer and intra observer agreement were quantified with the intraclass correlation coefficient (ICC) (Ebel 1951; Fleiss 1986) which has a range of 0 to 1 (perfect agreement). The ICC estimates the magnitude of true variation between subjects relative to the total variation in scores. The estimates for the variance are derived from the analysis of variance. As the inter-rater reliability assessment included different pairs of ratings for each subject, the variance was obtained from a one-way analysis of variance. However, when considering the intra-rater reliability, the variance estimates were derived from a two way analysis of variance, using subjects and time as the factors. Time was included to examine whether the sequence of evaluation systematically influenced the scores. For example, within a given pair, did the second test score tend to be higher than the first?

The reliability estimates and confidence limits for the ICC were performed for all subjects and separately for each longitudinal study. Reliability coefficients of .80 and above are generally considered high, but, when making decisions about individuals, more stringent criteria are recommended (Helmstadter 1964; Nunnally 1978).

3.3.3. Internal consistency

Methods and Analysis

The assessment of the third type of reliability, internal consistency was based on Balance Scale scores given by the independent evaluators separately within each longitudinal study.

The descriptive statistics for the items in the Balance Scale included the frequency distribution of scores for each of the five response categories, the mean score for each item and the item-to-total correlations. The magnitude and direction of the Pearson's correlation coefficents of each item with every other item in the Balance Scale were examined in a correlational matrix. The internal consistency was tested by Cronbach's Alpha (Cronbach 1951) for each evaluation time. The underlying assumption of this statistic is that each item is considered to be measuring the same common concept and thus the sum is likely to give a better estimate than any single item. The more the items covary relative to the sum of their variance the higher the Cronbach's Alpha. Standardized item alpha refers to the alpha value that would be obtained if all the items are standardized to have a variance of 1. In that case, Cronbach's Alpha can be based on the average correlation of items in a scale. Cronbach's Alpha is regarded as high if greater than .80. An item-to-total correlation shows the degree of association between each individual item and the total score of the other items in the scale. An item-to-total correlation is considered adequate if it is above 0.4.

To assess the consistency of the findings each aspect of the analysis was repeated at each evaluation point for the two study populations, the elderly residents and the patients with a diagnosis of stroke.

3.3. Summary

Several strategies were developed to further assess the measurement properties of the Balance Scale. Validity Study I examines the criterion validity of the Scale relative to clinical judgments of balance, use of mobility aids and the occurrence of subsequent falls in the elderly residents. Validity Study II assesses construct validity based on the hypothesized relationship between balance, functional status and motor performance in the early recovery of stroke patients. The ability of the Balance Scale to monitor changes in the status of stroke patients is assessed relative to the Barthel Index. In addition, the criterion validity of the Balance Scale is examined by comparing stroke patients by their place of residence at the time of each evaluation. Lastly, the Reliability Study assesses the inter and intra-rater reliability and internal consistency in elderly residents and stroke patients.

Chapter 4 RESULTS OF THE MEASUREMENT STUDY

4.0. Introduction

The results of the validation procedures follow the same sequence as the presentation of the methods. Validity Study I addresses the three objectives set forth within the study of elderly residents. Validity Study II reports the results for the stroke population. The last part presents the findings of the reliability study which used both elderly residents and stroke patients as subjects.

4.1. Validity Study I

In total, 115 residents of the home for the elderly volunteered to participate in the study. Of these, one withdrew from the study prior to the first evaluation and another almost immediately left the residence to live with her brother. Both have been excluded from the analysis, leaving 113 subjects at entry.

4.1.1. Characteristics of the subjects

Table 4.1 displays the sociodemographic and clinical characteristics of residents at entry to the study. The residents were predominantly female (82%), English speaking (90.3%) and well-educated with an average of 12.6 years of schooling. Their average age (83.5 years) was slightly below the overall mean of 86.6 years of all residents at the Griffith-McConnell Residence.

The medical diagnoses of subjects were classified according to system impairment. In this sample, the most common conditions were cardiovascular diseases (55.8%), hypertension (52.2%) and rheumatic diseases (43.4%). Not all systems were represented on the coding sheets. The most frequently occurring conditions in the "other" category were depression, hypothyroidism and anemia. Residents had a mean of 3.9 (SD 1.4) associated conditions for which they took, on average, 3.9 (SD 2.0) medications.

Overall, subjects were quite independent in the basic activities of daily living with a mean Barthel Index score of 98.3 (SD 4.2). The mean Balance Scale score was 46.8 (SD 6.6) and the average Mini-Mental State Examination score (Folstein et al 1975) was

TABLE 4.1

Sociodemographic and clinical characteristics of elderly residents at entry to the study (N=113)

Sociodemographic Characteristics	**Mean (SD)** Number (percent)	Clinical Characteristics	**Mean (SD)** Number (percent)
Age (years)	**83.5 (5.3)**	Medical Problems	
		Neurological	46 (40.7)
Sex		Cardiovascular	63 (55.8)
Male	20 (18.0)	Pulmonary	16 (14.2)
Female	93 (82.0)	Diabetes	9 (8.0)
		Peripheral Vascular	11 (9.7)
Marital Status		Rheumatic	49 (43.4)
Married	11 (9.7)	Visual	45 (39.8)
Never Married	29 (25.7)	Hypertension	59 (52.2)
Formerly Married	73 (64.6)	Gastrointestinal	26 (23.0)
		Genitourinary	9 (8.0)
Language		Neoplasm	16 (14.2)
French	10 (8.8)	Orthopaedic	28 (24.8)
English	102 (90.3)	Other	58 (51.3)
Other	1 (0.9)		
		Mean # Diagnoses	**3.9 (1.4)**
Education (years)	**12.6 (3.9)**		
		Mental Status (MMSE)	**27.9 (2.7)**
Usual Occupation			
Professional	45 (39.8)	Mobility Aids	
Clerical	31 (27.4)	None	49 (43.4
Sales	5 (4.4)	Cane outdoors	26 (23.0)
Service	3 (2.7)	Cane indoors	29 (25.7)
Transportation	2 (1.8)	Walker	9 (8.0)
Crafts	3 (2.7)		
Labourer	1 (0.9)	Medications	
Housewife	21 (18.6)	Major tranquilizers	6 (5.3)
No occupation	2 (1.8)	Sedatives	42 (37.2)
		Antiplatelet/coagulant	28 (24.8)
		Antidepressant	10 (8.8)
		Antihypertensive	51 (45.1)
		Cardiac	48 (2.5)
		Antiinflammatory	36 (31.9)
		Other	85 (75.2)
		Mean # Medications	**3.9 (2.0)**

within the normal range for this age group, 27.9 (SD 2.7) out of a possible 30 points. Only 9 residents scored below 24, indicating possible cognitive impairment.

Over half of subjects lived in the McConnell, one third in the Griffith (31.9%) and the remaining (12.4%) were in the Annex at the time of their first evaluation. These percentages are approximately in proportion to the numbers of residents living in each section, indicating recruitment was fairly balanced by site. In fact, the number of residents from each section that participated in the study corresponded to 44%, 35% and 34% of available beds in the McConnell, the Griffith and the Annex sections respectively. During the year of follow-up, certain subjects moved to a more dependent area permanently or temporarily due to illness but none were recruited from the infirmary or the special care unit.

The clinical and functional scores of subjects within each section were consistent with their expected level of physical activity and functional independence. There were gradients in scores for the Barthel Index, the Mini-Mental State, and the Balance Scale scores depending on where the resident lived within the Griffith-McConnell at the time of the baseline evaluation. The participants from the Annex had the lowest mean scores for each measure. As previously mentioned, the residents in the Annex receive compulsory supervision for bathing and have the option of assistance in other areas. The distinction between the level of functioning in basic activities of daily living was not so clear between the residents of the McConnell and the Griffith as both scored at the top of the Index.

4.1.2. Handling of missing data

Each follow-up evaluation point during the year has information missing for some subjects. No data were imputed for the missing values. Table 4.2 describes the study losses in relation to whether or not we were able to obtain the Balance Scale score. The table shows the time a permanent withdrawal occurred and, in the case of individuals missing occasional information, it lists the total number of subjects who did not perform a Balance Scale assessment at each evaluation point.

Five subjects missed one evaluation due to illness, travel, or unwillingness to

TABLE 4.2

Reasons for missed evaluations during the year

TABLE 4.2.1 Reasons for a permanent withdrawal from the study during the year (N=12)

	TIME OF WITHDRAWAL		
	Three-Month Evaluation	Six-Month Evaluation	Nine-Month Evaluation
Death	1	2	1
Stroke	0	1	0
Fracture of lower extremity	1	3	1
Voluntary refusal to continue	0	0	2*

TABLE 4.2.2 Reasons for missed Balance Scale evaluations during the year (N=9)**

	TIME OF MISSED EVALUATIONS		
	Three-Month Evaluation	Six-Month Evaluation	Nine-Month Evaluation
Voluntary refusal	4*	3	2
Illness	1	1	1
Vacation	0	0	1

* One of the residents missed the 3-month and withdrew from the study at the 9-month evaluation.
** Three residents missed more than one evaluation.

comply with an assessment at the prescribed interval. Three subjects missed more than one balance evaluation, but of these, two completed the corresponding Barthel Index, self-rating and the fall interim report. One subject missed only the 3-month Balance evaluation, completed all aspects of the 6-month and then withdrew from the study at 9 months. He is included both in the temporarily missed evaluations and as one of the two voluntary withdrawals at 9 months.

The permanent losses to the study included four individuals who died, five with lower extremity fractures, and one who suffered a stroke. The latter six residents were excluded from subsequent analyses because of the restrictions placed on their mobility and ability to bear weight through their legs. They required a period of hospitalization for medical or rehabilitation services. All but one returned to the residence within the year of follow-up.

A resident who fractured her arm is considered to have had a temporary illness at the 6-month evaluation because, in contrast, to the residents with lower extremity fractures, she did not leave the residence and had no restriction placed on her ability to bear weight on her lower extremities. Another resident with an upper extremity fracture completed all four evaluations. Similarly, a resident fractured her lower extremity following the 9-month Balance Scale assessment.

All reported falls were consistent with the definition suggested by the Kellogg International Work Group (1987): "A fall is an event which results in a person coming to rest inadvertently on the ground or other lower level and other than a consequence of sustaining a severe blow, loss of consciousness, sudden paralysis, as in stroke, or an epileptic seizure". The occurrence of falls for all subjects was ascertained through interviews with the residents, staff reports and official records. Thus, although complete balance evaluations were available for only 93 residents, the analysis examining the relationship between the Balance Scale scores and the occurrence of falls within the year includes all subjects entered into the study. The number of subjects in the analysis are contained within each table or figure.

4.1.3 Relationship of Balance Scale scores and the type of mobility aid (criterion validity)

The relatively high level of functioning of the residents is further reflected in the use of walking aids. Seventy-five residents did not use any mobility aids to walk indoors but, of these, 26 residents used a cane outdoors. Nine residents required a walker and 29 routinely used a cane.

The first objective was to determine if Balance Scale scores could discriminate between these groups using mobility aids that offer different amounts of support. As demonstrated in Table 4.3 mean Balance Scale scores showed a gradient from the high to low for the four groups: no aids (49.5), cane outdoors (48.3), cane (45.3) and walker (33.1). Despite an overlap in the distributions, the difference in means was greater than expected by chance alone when tested in a one-way analysis of variance. In addition, polynomial contrasts confirmed the linear relationship ($p<.0001$) observed in the gradient of mean scores.

The initial inspection indicated that the relationship between balance and use of mobility aids differed according to the MMSE scores of the residents. Mental status can influence a subject's self-perception of the need for extra support and whether he or she remembers to use an appropriate aid. The mean Balance Scale scores for persons using no walking aids (38 SD 6.7) was notably lower in the five subjects with MMSE scores below 24 than those with MMSE scores of 24 and above (50.8 SD 3.7).

Other factors potentially influencing the use of mobility aids were examined and considered as covariates in an analysis of variance. This analysis was restricted to the 104 subjects with MMSE scores of 24 and above. The inclusion of Barthel mobility scores, age and the presence of an orthopaedic diagnosis or previous fracture as covariates did not result in marked changes in the adjusted means, and the observed differences in mean Balance Scale scores remained greater than expected by chance alone. A neurological or rheumatological diagnosis was not helpful in explaining more of the variation in Balance Scale scores among subjects grouped by the level of walking aid.

The strength of the relationship between Balance Scale scores and the use of

TABLE 4.3

Balance Scale means and 95% confidence intervals for the mean of the elderly residents grouped by use of walking aids (N=113)

	BALANCE SCALE SCORES		
	N	Mean	(95% CI)
No Walking Aids	49	49.5	(47.9 - 51.1)
Cane Outdoors Only	26	48.3	(47.0 - 49.6)
Cane	29	45.3	(44.0 - 46.6)
Walker	9	33.1	(26.7 - 39.6)

$F_{df\ 3,109}$ 29.6
Test for linear effect
$p<.0001$

walking aids was also examined by the Spearman Rank Order Correlation Coefficient. The magnitude of the coefficient was moderately strong (rho -.56) and in the expected direction, namely, the lower the Balance Scale score the greater the need for support. When the 9 subjects with suspected cognitive impairment were omitted from the analysis, the correlation coefficient assessing the strength of the relationship between the Balance Scale scores and the use of mobility aids increased to rho= -.67.

4.1.4. The relationship between Balance Scale scores and global ratings (concurrent criterion validity)

The second objective was to determine the degree to which Balance Scale scores were related to concurrent global ratings of good, fair or poor balance given by the senior matrons or nurses familiar with the residents. Figure 4.1 demonstrates the distribution of the Balance Scale scores within each category at each evaluation point. The percentage of residents in the good, fair and poor categories were on average 55%, 35% and 10% respectively.

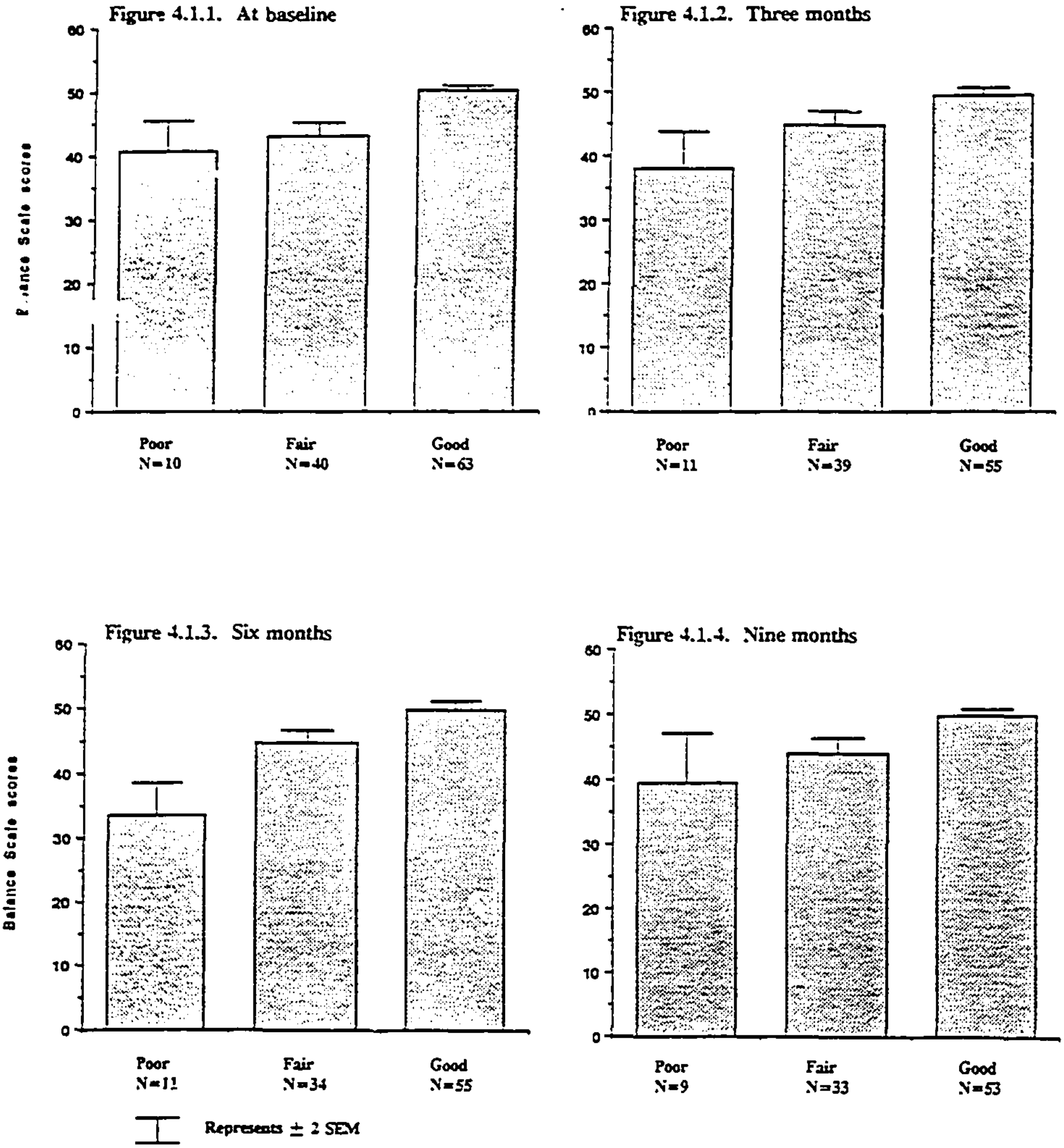
The primary analysis tested the strength of the relationship between Balance Scale scores and global ratings. The Spearman's Rank Order correlation coefficients between the two assessments at each of the four evaluation points in the study ranged from .50 to .62, indicating a moderately strong positive relationship.

At each evaluation, the residents judged to have good balance at that point in time had mean Balance Scale scores that remained at a similar level (49.5 to 50.2), as did those judged as fair (43.1 to 44.7). In contrast, the individuals rated as poor at that evaluation displayed greater variability and had mean scores that ranged from 33.5 to 40.8. When assessed by a one-way analysis of variance, the difference in means between the three categories at each evaluation point was greater than expected by chance alone. There was a linear gradient ($p<.0001$) of the mean scores. In addition, the difference in Balance Scale means between the good, fair and poor categories remained after using the Barthel Index scores and age as covariates in the model. Details of the mean scores for each category of rating are included in Appendix 4.1.

Spearman's Rank Order correlations between the subject's self-rating of his or her

FIGURE 4.1

Balance Scale scores for good, fair and poor global judgments at each evaluation of the elderly residents



ability and the Balance Scale scores were somewhat lower (.43 to .49) but still of moderate magnitude and in the expected direction.

4.1.5. Relationship of Balance Scale scores and falls (predictive validity)

The third objective was to assess the relationship between the Balance Scale scores and falls in the year of follow-up. Falls were ascertained through self-reports of the residents, verbal reports by staff, official incident reports and chart information. Sixty (53.1%) study participants did not fall, 24 (21.2%) participants fell once and 29 (25.7%) had two or more falls in the year. There is a gradient in mean initial balance scores of 48.8 (SD 4.4), 46.8 (SD 7.6), 42.8 (SD 7.7) for non-fallers, one-time and multiple-time fallers respectively. However, the association with fall status (0, 1 and 2 or more falls) remains quite low (rho -.36; $p<.01$), suggesting that other factors contribute to the occurrence of falls.

The clinical and functional characteristics of residents were compared across three groups having none, one and multiple falls in the year of follow-up. A one-way analysis of variance showed that the observed differences in means were greater than expected by chance for Balance, Barthel mobility sub-scale and the average number of medications. The gradients in the mean scores of the Barthel Index and the Balance Scale scores were in the expected direction, namely, the non-fallers had the highest scores and the multiple fallers has the lowest. The mean number of medications and diagnoses did not show the same pattern. Non-fallers took the greatest number of medications (mean 4.3, SD 1.8). Furthermore, residents who fell once had the lowest mean for medications (3.2, SD 1.8) and number of diagnoses (3.4, SD 1.4). Details on the functional and clinical characteristics are included in Appendix 4.2.

The distribution of cardiovascular disease and visual deficits showed marked differences across the three groups (chi square $p < .05$). Neither showed a gradient effect from non-fallers to multiple fallers. Individuals who had one fall reported less cardiovascular disease compared to the other groups whereas multiple time fallers had higher frequency of visual deficits. This finding reinforced the decision to consider visual deficits as a risk factor for falls.

The mean Balance Scale (range 46.9-47.7) and Barthel Index (range 97.5-98.5) scores remained stable throughout the year for subjects with complete information. A repeated measures analysis examining within subject changes for the four evaluations during the 9-month period did not demonstrate a difference in either the Balance Scale or Barthel Index scores. The relative stability of the scores during the year justified the use of the initial Balance Scale score as a risk factor for falls throughout the year.

The relative risk of falling during the year was 1.5 (95% CI 1.0-2.3) for the 31 subjects scoring below 45 at the initial evaluation compared to those with higher scores. Forty-five was chosen as a cut-off based on verbal reports from clinicians familiar with the Balance Scale. The relative risk was higher, 2.7 (95% CI 1.5-4.9), when multiple falls was the outcome of interest. Thus, subjects with scores below 45 are 2.7 times as likely to fall more than once compared to residents with scores of 45 and above.

Eight residents suffered a fall-related fracture during the study. The initial Balance Scale scores of the subjects with subsequent fractures were on average 38.9 (SD 12.2).

Logistic regression was used to determine which variables were predictors of falling at least once during the year. Age, cognitive impairment, Barthel Index and its mobility sub-scale, visual impairment, total number of diagnoses as well as orthopaedic and rheumatic diagnoses were considered as potential variables in the models. Barthel Index and Barthel mobility sub-scale scores were not included in the same models due to their high inter-correlation. Balance Scale scores showed low to moderate correlations with these variables (.03 to .57). The other variables were poorly intercorrelated. Balance Scale scores and the variable relating to the fact that the patient had fallen during the three months prior to the onset of the study were the only variables to enter the regression model at the 0.05 level (Table 4.4.1).

Given the apparent differences between single and multiple time fallers, the analysis was repeated using multiple falls as the outcome. The results are presented in Table 4.4.2. Balance Scale scores and history of falling remained useful predictors of falls. Additionally, the presence of visual deficits emerged as a risk factor for multiple falls. The other potential variables did not contribute additional information about the risk of falling in this population.

TABLE 4.4

Logistic regression models predicting falls in the year (N=113)

TABLE 4.4.1 **Logistic regression model predicting at least one fall compared to no falls among the elderly residents**

Variables	Beta Coefficient	Standard Error	Significance Level	Adjusted Odds Ratio	95% CI
*Initial Balance Scale Scores	-.11	.04	.005	.90	.83 to .97
History of falls in past 3 months	1.62	.70	.021	5.04	1.28 to 19.90
Constant	4.86	1.87	.010		

TABLE 4.4.2 **Logistic regression model predicting multiple falls compared to single or no falls**

Variables	Beta Coefficient	Standard Error	Significance Level	Adjusted Odds Ratio	95% CI
*Initial Balance Scale Scores	-.11	.04	.006	.90	.83 to .97
History of falls in past 3 months	1.75	.64	.007	5.74	1.63 to 20.25
Visual Deficits	1.03	.50	.039	2.80	1.05 to 7.44
Constant	3.83	1.75	.03		

* In each analysis, the odds ratio refers to the change in the odds of falling with a change of one point in the Balance Scale given that all other factors remain the same.

The negative beta coefficient shown in Table 4.4.2 for the Balance Scale indicates a protective effect for individuals with higher Balance Scale scores (adjusted-odds ratio .90). In contrast, having a visual deficit (adjusted odds ratio 2.8) or having experienced a recent fall (adjusted odds ratio 5.8) increase the odds of falling more than once. When comparing the magnitude of the odds ratios of the three variables it is important to consider that history of falls and presence of visual deficits are dichotomous variables, whereas the Balance Scale scores are continuous. Therefore the odds ratio for the Balance Scale represents a change in the odds of falling with each single point of the Scale. Hence, the odds of falling increases threefold (95% CI 1.4-6.7) with every drop of 10 points on the Balance Scale, if all other factors remain the same.

The consistency of the contribution of impaired balance to the occurrence of falls is further illustrated by comparing the mean Balance Scale of subjects who either did not fall, had a single fall or had two or more falls in each 3-month interval. The intervals were considered separately, with subjects potentially classified differently within each interval. As shown in Table 4.5, the gradient in mean scores found in the initial Balance Scale scores remained fairly consistent within each interval.

The relationship between Balance Scale scores and falls at each interval, the elevated relative risk of falling associated with a score below 45 and the contribution of the initial Balance Scale score to the occurrence of falls in the logistic regression model are findings that support the predictive validity of the Balance Scale.

Summary of Validity Study I

Balance Scale scores were assessed relative to three external criteria: use of mobility aids, clinical judgments of balance and occurrence of future falls. The results of each analysis provided support for the validity of the measure. Balance Scale scores showed a linear gradient in scores from high to low for the four groups: no aids (49.5), cane outdoors, (48.3), cane (45.3) and walker (33.1). They showed a moderately strong positive relationship (rho .50 to .62) with clinical judgments at each of the four evaluation points. Lastly, the Balance Scale scores showed an association with the occurrence of falls in several analyses: Spearman's rho correlation (-.36), relative risk

TABLE 4.5

Means and standard deviations (SD) of the Balance Scale Scores of the elderly residents grouped by whether or not they fell in the subsequent interval

INTERVAL	NO FALLS within interval Mean (SD)	SINGLE FALL within interval Mean (SD)	MULTIPLE FALLS within interval Mean (SD)
Baseline → 3 months	47.2 (6.2) N=107	0	38.8 (9.8) N=5
3 months → 6 months	47.5 (6.6) N=81	43.3 (7.7) N=17	43.3 (7.6) N=8
6 months → 9 months	47.7 (6.5) N=78	42.4 (9.1) N=19	33.3 (7.7) N=4
9 months → 1 year	47.3 (6.5) N=79	46.7 (5.4) N=15	36.7 (11.6) N=3
Baseline → 1 year	48.8 (4.5) N=60	46.8 (7.6) N=24	42.8 (7.7) N=29

The number of falls by an individual subject is counted separately for each interval. A given individual can appear in a different column at each interval.

elevated for individuals with Balance Scale scores below 45 (RR 2.7 for multiple falls; RR 1.5 for at least one fall), and Balance Scale scores were a useful predictor in the logistic regression model (adjusted odds ratio .90).

4.2 Validity Study II

Three specific objectives were formulated to assess construct validity, responsiveness and concurrent criterion validity of the Balance Scale in acute stroke patients. Seventy stroke patients met the entry criteria and consented to join the study. Patients were recruited in the general hospital within two weeks of stroke onset and re-assessed at 4, 6 and 12 weeks.

4.2.1 Reasons for non-response or exclusion

In total, 238 subjects were screened, 89 considered eligible and 70 enrolled in the study. Of the non-participants, 120 and 48 came from the Royal Victoria-MNI Complex and Montreal General Hospital, respectively. The reasons for exclusion are listed in Table 4.6. The two most common factors making patients ineligible were lack of motor or sensory deficits and an unstable medical condition. Information concerning the eligibility of 15% of the subjects was unclear due to communication difficulties with the subject and the absence of responsible family member or friend.

4.2.2 Characteristics of the subjects

Table 4.7 displays the sociodemographic and medical characteristics of the 70 stroke patients at entry to the study. The mean age of the subjects was 71.6 years. The majority (95.7%) lived at home prior to the onset of the stroke, but only 5.7% were working full-time. The proportion of males (51.4%) to females was approximately equal, as was the side of impairment. Patients had a mean of 2.6 (SD 1.4) associated medical conditions for which they took on average 3.8 (SD 1.9) medications.

4.2.3 Handling of missing data

The analysis of the data for stroke patients is based on information about 60

TABLE 4.6

Reasons for exclusion of stroke patients by hospital of admission

Reasons	Royal Victoria MNI Complex N=120	Montreal General Hospital N=48
No motor/sensory problems	27	3
Medically unstable	19	13
Information unavailable	19	7
Identified more than 2 weeks post onset of stroke	9	2
Speaks neither English nor French	7	7
No Medicare	3	1
Resides outside boundary	5	1
Diagnosis other than stroke	1	0
Age less than 40 years	4	0
Died before evaluation	11	1
Co-morbidity prohibits rehabilitation	7	2
Patient refused participation	6	4
Consent refused/family	2	7

TABLE 4.7

Sociodemographic and clinical characteristics of the stroke patients at enrolment (N=70)

Sociodemographic Characteristics	Mean (SD) Number (percent)	Clinical Characteristics	Mean (SD) Number (percent)
Age (years)	**71.6 (10.1)**	Side of weakness	
		Right	38 (54.3)
Sex		Left	32 (47.7)
Male	36 (51.4)		
Female	34 (48.6)	Comorbidity	
		Neurological	11 (15.7)
Marital Status		Cardiovascular	40 (57.1)
Married	41 (58.6)	Pulmonary	8 (11.4)
Never Married	9 (12.9)	Diabetes	22 (32.9)
Formerly Married	20 (28.6)	Peripheral Vascular	8 (11.4)
		Rheumatic	4 (5.7)
Language		Visual	5 (7.1)
French	22 (31.4)	Hypertension	46 (65.7)
English	28 (40.0)	Renal	4 (5.7)
Other	20 (28.6)	Gastrointestinal	11 (15.7)
		Genitourinary	6 (8.6)
Usual Living		Neoplasm	1 (1.4)
Arrangements		Orthopaedic	4 (5.7)
Home Alone	21 (30.0)	Other	9 (12.9)
Family	46 (65.7)		
Residence	2 (2.9)	Mean # Co-morbid	**2.6(1.4)**
Institution	1 (1.4)	Conditions	
Employment Status		Medications	
Full-time	4 (5.7)	Sedatives	9 (12.9)
Retired	49 (70.0)	Antiplatelet/coagulant	54 (77.1)
Unemployed	2 (2.9)	Antidepressant	2 (2.90
Housewife	13 (18.6)	Antihypertensive	36 (51.4)
		Cardiac	35 (50.0)
Usual Occupation		Antiinflammatory	5 (7.1)
Professional	11 (15.7)	Other	46 (65.7)
Clerical	6 (8.6)		
Sales	4 (5.7)	Mean # Medications	**3.8 (1.9)**
Service	5 (7.1)		
Transportation	6 (8.6)		
Production	8 (11.4)		
Labourer	11 (15.7)		
Housewife	19 (27.1)		
Education (years)	**8.6 (3.5)**		

subjects who were followed for 3 months. When compared to those who completed all assessments, the ten patients who did not complete the study were on average older (76.6 years), had lower scores on the Balance Scale (7.2), the Arm and Leg sub-scales of the Fugl-Meyer (26.9), and the Barthel Index (13.5). The latter measure was the only one to show a difference that was greater than expected by chance alone ($p < .05$). The most common reason for a permanent loss from the study was death. Eight patients died during the 3-month follow-up period. Two did not complete the study because they were no longer permitted to bear weight on their lower extremities; one developed gangrene and another fractured his hip. Further details on the sociodemographic and medical characteristics of the subjects who did not complete the study are included in Appendix 4.3.

Values were imputed for five subjects with missing information at one evaluation. The imputed values were selected by the study team, based on the patient's performance at other assessment points. The reasons for the missing information varied. One man consented, was entered into the study but had an emergency transfer to the intensive care unit before the first evaluation could be done. At the time of his scheduled 4-week evaluation, he was on the ward and willing to continue in the study. Given the paucity of stroke patients, it was decided to retain him. Balance Scale and Barthel Index scores were imputed as zero, the same score as on the three subsequent assessments. The Fugl-Meyer motor performance for the arm and leg were imputed as 2 for each, whereas the Fugl-Meyer balance component was scored as 1. These values correspond to the scores for both the 4 and 6-week evaluations.

One other patient had a missing value for the initial Balance score. Her Fugl-Meyer Scale and Barthel assessments were complete. The balance evaluation was not completed due to a scheduling error that could not be corrected without violating the time constraints of the protocol. All subsequent evaluations were completed. A score of 22 was imputed for the Balance Scale based on the patient's other assessments and the scores of patients with similar profiles.

At six weeks, one man was assessed on the Barthel Index and the Balance Scale but refused to cooperate with the Fugl-Meyer assessment. During this period, he was still in the general hospital and not participating in any rehabilitation program. The imputed scores for the Fugl-Meyer arm, leg, and balance sections were 10, 13 and 4. The scores are intermediate to the values he received at four and 12 weeks. Overall, this man showed little recovery in the three months of the study.

The other two patients refused to participate in the final evaluation. They both required help in all activities and were still in the general hospital three months after their stroke. Given the lack of documented progress, scores equivalent to the 6-week evaluation were imputed for the 2 subjects for the 12-week assessment.

The missing values reflected the reality of clinical investigations conducted over time in several settings. Fortunately, they were relatively few in number. Values were imputed only for subjects whose status was observed at the end of three months. The imputed values were individualized based on available information and careful deliberation by the study team. It is unlikely that use of the imputed values altered any results obtained.

4.2.4. Relationship between the Balance Scale scores and functional status and motor performance (construct validity)

The first objective was to assess the degree of association between the Balance Scale and the motor performance and functional status scores at various points in the recovery of stroke patients. Table 4.8 presents the Product Moment Correlations between the Balance Scale, the Barthel Index and the Fugl-Meyer Scale and their respective subscales at each evaluation point for the 60 patients who completed the study. Correlations are all moderately-strong to strong and in the expected direction.

As indicated in the methods, it is clinically recognized that stroke patients fall into three categories: those who initially have a low level of function and show little recovery, others who have a high level of function initially that is maintained, and those who start low and make significant improvement in function. Patients with a Barthel Index score below 40 at entry and at the end of three months were classified as the "low functioning group". Patients entering the study with a score of 60 and maintaining that level were included in the "high functioning group". The remaining 32 subjects were placed in the

TABLE 4.8

Product-Moment correlations of the Balance Scale with the Barthel Index and Fugl-Meyer Scores for stroke patients at each evaluation point (N=60)*

	BALANCE SCALE			
	Initial Evaluation	**4-Week Evaluation**	**6-Week Evaluation**	**12-Week Evaluation**
Barthel Index	.90	.87	.90	.93
ADL Subscale	.81	.80	.82	.86
Mobility Subscale	.92	.85	.92	.94
Fugl-Meyer Scale	.70	.77	.77	.82
Arm Subscale	.62	.69	.69	.76
Leg Subscale	.71	.80	.76	.79
Balance Subscale	.84	.87	.88	.94

* All subjects who finished the study

group showing major improvement in functional ability. Table 4.9 shows mean scores on the three clinical measures for all subjects and by clinical sub-groups.

Figures 4.2.1-4 illustrate the covariation of the Balance Scale, the Barthel Index and the Fugl-Meyer scale measuring the motor performance of the arm and leg at the four evaluation points. The changes in the mean Balance Scale scores parallel those of the Barthel Index both in the total sample and in the three sub-groups. The relationship among the three measures remains consistent within the three sub-groups.

4.2.5. Monitoring the status of stroke patients: comparison of the Balance Scale scores to Barthel Index scores (responsiveness)

The similar pattern of improvement in mean scores of the Balance Scale and the Barthel Index over the 12-week period suggests that the Balance Scale is sensitive to changes in status. To quantitatively assess this property, a three-step procedure based on the results of the 60 subjects who completed the study, was used.

First, a repeated measures analysis of variance using polynomial contrasts was performed to examine within subject changes in the Barthel Index over the 12 week period. The results indicated a strong linear effect. Second, a similar analysis determined that the Balance Scale also had a strong linear effect showing improvement in scores over the 12-week period.

The third step examined whether the magnitude of the changes in the Balance Scale corresponded to those in the Barthel Index for the same patients. As shown in Figure 4.3, the mean differences between the Balance Scale and the Barthel Index remained relatively constant over the 12-week period, suggesting it is at least as good at monitoring the status of stroke patients as the Barthel Index. This observation was confirmed by a repeated measures analysis of variance examining the within subject changes at each evaluation and using the difference between the Barthel Index and the Balance Scale scores for each subject as the dependent variable.

In contrast, Figure 4.3 shows that the difference scores between the Barthel Index and the Fugl-Meyer Arm and Leg Sub-scales follow a linear pattern. The values for the first three evaluations below the zero line indicate that the combined Fugl-Meyer Arm

TABLE 4.9

Mean Barthel Index, Balance Scale and Fugl-Meyer Arm & Leg Sub-Scale for each clinical performance sub-group (N=60)*

	All Subjects With Complete Scores N=60		Low Functioning Group N=15		Group With Greatest Functional Change N=32		High Functioning Group N=13	
	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)
Barthel Index (0-100)								
Initial	33.8	(30.5)	4.0	(5.9)	28.4	(16.7)	81.5	(13.2)
4 Weeks	48.5	(34.1)	8.6	(10.6)	51.1	(26.0)	88.0	(11.5)
6 Weeks	55.5	(35.1)	10.8	(11.5)	61.0	(25.7)	93.4	(8.3)
12 Weeks	64.4	(34.9)	12.1	(12.3)	76.8	(18.6)	94.0	(11.3)
Balance (0-100)								
Initial	27.7	(29.9)	1.7	(2.3)	21.3	(18.9)	73.6	(14.4)
4 Weeks	42.7	(35.8)	3.6	(3.5)	45.0	(29.9)	82.3	(18.8)
6 Weeks	47.4	(36.5)	4.5	(6.4)	52.4	(29.8)	84.5	(19.5)
12 Weeks	56.7	(36.5)	7.9	(12.5)	66.7	(26.7)	88.2	(14.0)
Fugl-Meyer Arm and Leg (0-100)								
Initial	47.3	(35.2)	16.6	(19.5)	48.8	(33.5)	79.2	(21.2)
4 Weeks	56.7	(35.0)	19.3	(21.2)	64.0	(31.5)	81.8	(19.0)
6 Weeks	61.6	(33.0)	24.3	(21.3)	69.7	(28.3)	84.6	(17.0)
12 Weeks	64.6	(32.7)	26.0	(23.6)	73.5	(26.6)	87.1	(12.6)

* All subjects who finished the study

FIGURE 4.2

Mean Balance Scale, Barthel Index and Fugl-Meyer Arm and Leg Sub-scale scores of the stroke patients at each evaluation (N=60)

Figure 4.2.1.
All subjects (N=60)

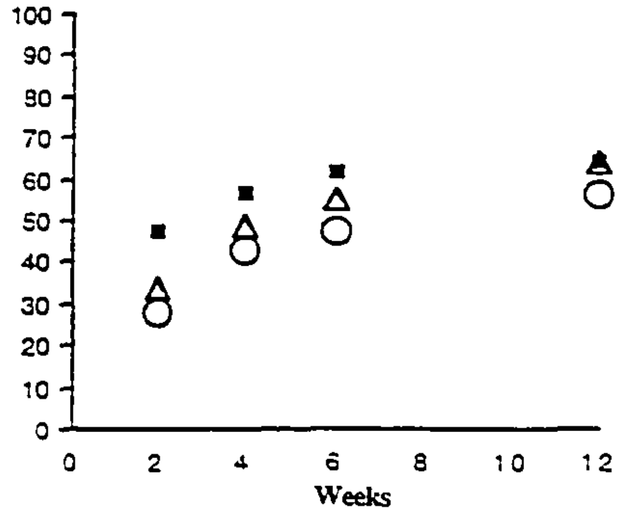


Figure 4.2.2.
High functioning group (N=13)

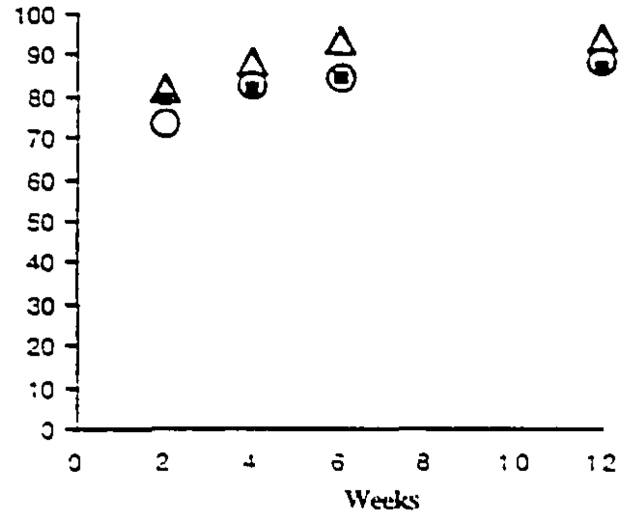


Figure 4.2.3.
Low functioning group (N=15)

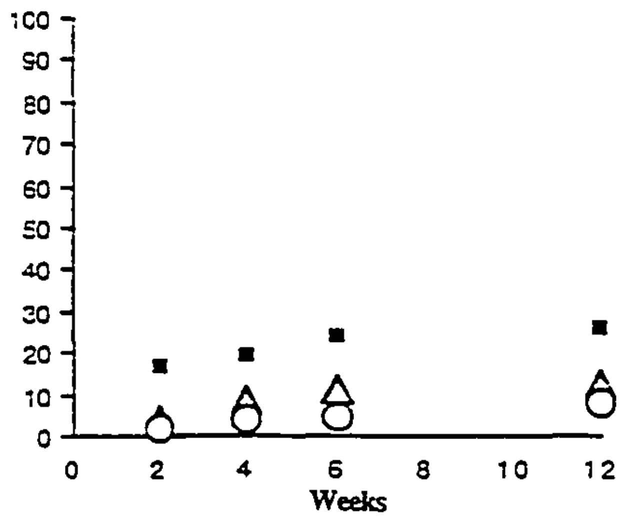
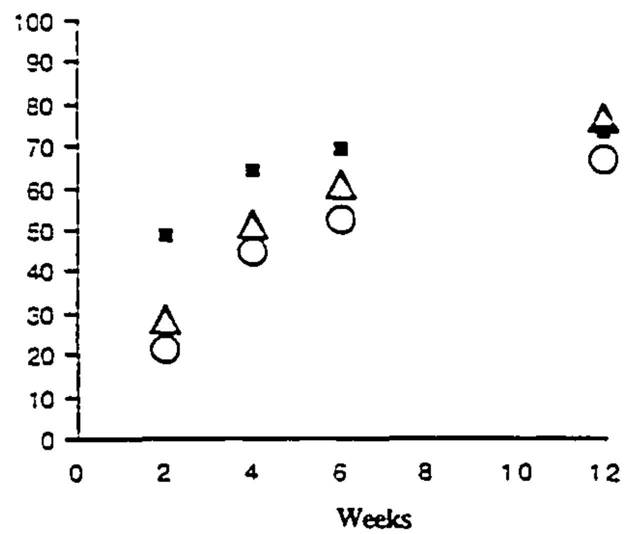


Figure 4.2.4.
Group with greatest functional gains (N=32)



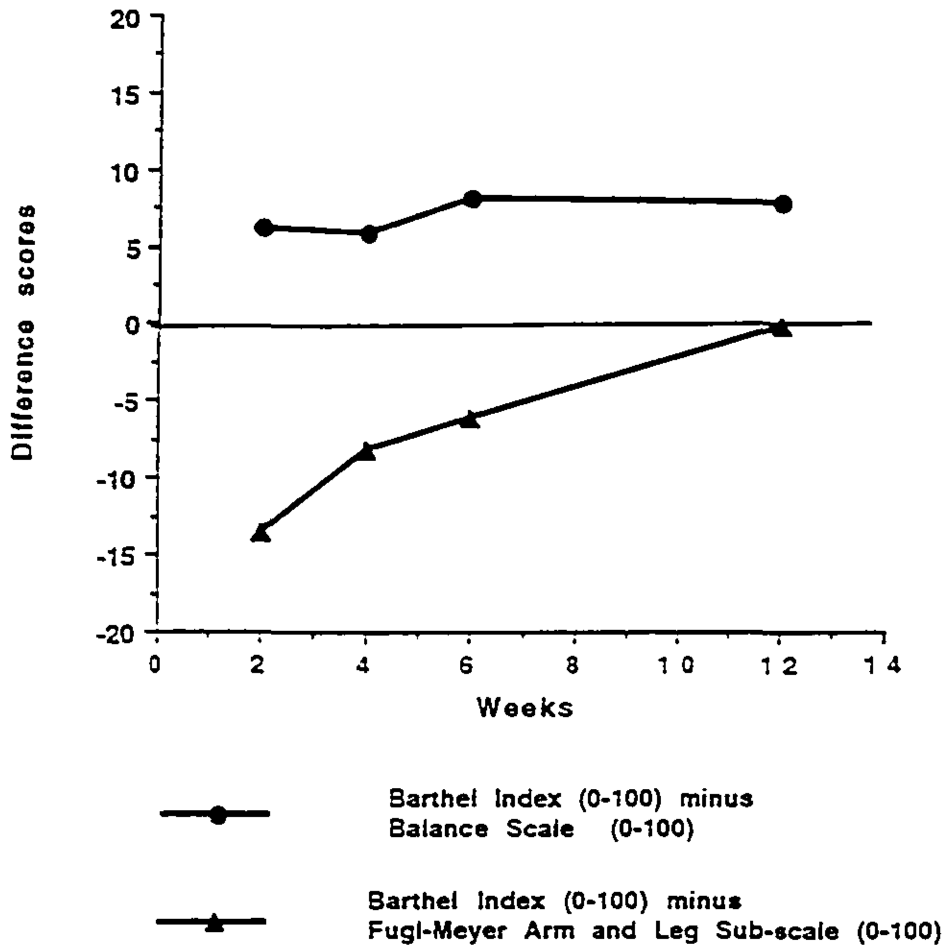
△ Barthel Index (0-100)

○ Balance Scale (0-100)

■ Fugl-Meyer Arm and Leg Sub-scales (0-100)

FIGURE 4.3

Mean differences between the Barthel Index and the Balance Scale scores and the Barthel Index and the Fugl-Meyer Scale scores for stroke patients at each evaluation (N=60)



and Leg sub-scale scores were higher than the Barthel Index scores. The fourth evaluation bisects the line, showing no difference in scores at that point. This finding is supported by Figures 4.2.1-4 showing the higher mean scores initially and the flattening of the slope as the rate of change slows and approximates the Barthel Index scores. The repeated measures analysis also confirmed the linear effect of the change scores over time, suggesting that the Fugl-Meyer Arm and Leg Sub-scales are not as sensitive in detecting change as the Barthel Index.

Similarly, the Balance sub-section of the Fugl-Meyer showed that it was able to detect a linear effect for the within person changes over time, but that the magnitude of change over the study period was less then that seen in the Barthel Index.

Each of the performance sub-groups showed patterns and magnitude of change that paralleled those of the Barthel Index when assessed with the same three-step procedure as above. In addition, the comparison among the clinically defined sub-groups was consistent with the expectation of a responsive measure. A responsive measure should show the greatest increase in scores in the group demonstrating the most functional improvement and should remain relatively stable in the group expected to change little (low functioning group). Both Table 4.9 and Figure 4.2 are consistent with this expectation.

Additional evidence of the responsiveness of the Balance Scale is that each item of the scale showed a higher mean score at each subsequent evaluation. Only two items did not show improvement from 4 to 6 weeks. Table 4.10 shows the trend of improving mean scores over time. The magnitude of the item means also illustrates the hierarchy or degree of difficulty of the items in the scale.

4.2.6. Relationship between the Balance Scale scores and the place of residence at each follow-up evaluation (concurrent criterion validity)

The mean Balance Scale scores at 4, 6 and 12 weeks were compared across groups defined by where they lived at the time of the follow-up evaluation. As shown in Table 4.11, there was a gradient in mean scores from high to low for the subjects living at home, in a rehabilitation setting, and those still in the general hospital at each

TABLE 4.10

Mean scores for each item of the Balance Scale for stroke patients at each evaluation (N=57)*

Scale Item	Two Weeks Mean (SD)	Four Weeks Mean (SD)	Six Weeks Mean (SD)	Twelve Weeks Mean (SD)
Sit to Stand	1.21 (1.46)	2.04 (1.59)	2.35 (1.64)	2.61 (1.61)
Standing	1.53 (1.62)	2.09 (1.84)	2.39 (1.74)	2.77 (1.72)
Sitting	2.63 (1.58)	3.21 (1.36)	3.35 (1.33)	3.54 (1.07)
Stand to Sit	1.39 (1.62)	2.02 (1.67)	2.26 (1.72)	2.61 (1.58)
Transfer	1.39 (1.42)	2.16 (1.42)	2.32 (1.57)	2.71 (1.37)
Stand Eyes Closed	1.49 (1.66)	2.04 (1.87)	2.30 (1.80)	2.70 (1.75)
Stand Feet Together	0.93 (1.41)	1.54 (1.75)	1.86 (1.83)	2.25 (1.78)
Arm Reaching	1.07 (1.45)	1.60 (1.57)	1.73 (1.59)	2.07 (1.55)
Object Pick Up	1.11 (1.63)	1.84 (1.79)	2.01 (1.81)	2.51 (1.73)
Twisting	1.21 (1.59)	1.91 (1.81)	1.91 (1.75)	2.35 (1.71)
Turn 360°	0.72 (1.22)	1.26 (1.61)	1.47 (1.75)	2.02 (1.83)
Step on stool	0.40 (0.88)	1.00 (1.48)	1.21 (1.63)	1.60 (1.77)
Tandem Standing	0.51 (0.95)	1.12 (1.36)	1.09 (1.48)	1.54 (1.46)
One Leg Standing	0.32 (0.66)	0.82 (1.23)	0.95 (1.39)	1.35 (1.58)

* Only patients with complete Balance Scale scores were used in this analysis. Item values were not imputed unless total Balance Scale score was zero.

TABLE 4.11

Means and 95% confidence intervals for the mean of the Balance Scale scores of the stroke patients grouped by the location of each follow-up evaluation (N=60)

	N	At 4 Weeks Mean (95% CI)	N	At 6 Weeks Mean (95% CI)	N	At 12 Weeks Mean (95% CI)
General Hospital	47	19.5 (13.8 - 25.2)	30	16.1 (9.0 - 23.2)	14	8.1 (1.8 - 14.5)
Rehabilitation Centre	4	27.3 (1.0 - 53.5)	13	32.9 (21.7 - 44.2)	20	31.1 (21.6 - 40.5)
Home/Community	9	45.3 (38.8 - 51.9)	17	40.1 (32.7 - 47.4)	26	45.0 (40.6 - 49.3)
		$F_{df\ 2,57}$ 7.76 $p<.001$		$F_{df\ 2,57}$ 11.05 $p<.001$		$F_{df\ 2,57}$ 28.67 $p<.001$

assessment. As expected, the number of patients still in the general hospital became smaller and their mean score became lower at each subsequent evaluation. At 12 weeks, the average Balance Scale score of the remaining 14 patients was 8.1, reflecting the fact that they are the most severely involved. Conversely, the number of subjects discharged to the community increased progressively and their mean Balance Scale score remained 40 or above at each subsequent evaluation. The mean for the patients at the Rehabilitation centres was intermediate, varying from 27.3 to 32.9.

A one-way analysis of variance performed at each evaluation demonstrated that the observed mean differences between subjects living at home, in a rehabilitation centre and in a general hospital were greater than expected by chance alone. This ability of the Balance Scale to discriminate between the mean Balance Scale scores of the groups provides evidence of concurrent criterion validity.

Summary of Validity Study II

The correlations between the Balance Scale and measures of functional status and motor performance were high (.70 -.93) at each evaluation. Moreover, the changes in the Balance Scale scores mirrored those of the Barthel Index in each of the clinical sub-groups and their magnitude was equivalent to the changes detected by the Barthel Index over the same period. Lastly, there was a linear gradient in mean Balance Scale scores for patients discharged home (range 45.3-45.0), discharged to a rehabilitation hospital (range 27.3-32.9) and still in the general hospital (19.5-8.1) for the follow-up evaluations.

4.3. Reliability Study

Three objectives were formulated to assess reliability. The first two examined the inter and intra-rater reliability of the Balance Scale in groups of elderly residents and stroke patients. The third objective was to assess the internal consistency of the Balance Scale. This analysis was performed on all available subjects at each evaluation point within both study populations separately.

4.3.1 Characteristics of the subjects participating in the reliability study

The 31 elderly residents and the 36 stroke patients who participated in the inter and intra-rater reliability study were similar in terms of sociodemographic and baseline clinical characteristics to the total sample from which they were chosen. Details of the comparison are included in Appendix 4.4 and 4.5.

4.3.2 Inter-rater reliability

In total, 32 individual raters (1 nurse, 21 physical therapists, 6 occupational therapists, 2 senior matrons and 2 physical therapy students) were used to rate 35 stroke patients and 28 elderly residents in the inter-rater reliability study. Each patient was evaluated twice, by random pairs of raters. The Balance Scale scores for all the subjects covered the entire range (0-56) of the scale and had a mean of 37.1 (SD 17.2), averaged over both ratings.

Figure 4.4 shows the pairs of ratings for each of the 28 elderly residents. The scores ranged from 25 to 55. The paired ratings appear consistent with most ratings within a few points of each other. The worst case showed a difference of 8 points.

Figure 4.5 presents the paired scores for the 35 stroke patients whose scores encompassed the entire range of the Balance Scale (0-56). The scores demonstrate generally good agreement, but there are occasional differences in certain subjects.

The intraclass correlation coefficient (ICC) was used to quantify the agreement between the raters. This statistic estimates the true variance between subjects relative to the total observed variance in scores. Variance estimates are obtained from the one-way analysis of variance. Overall, when all subjects were included, the ICC was .98 (95% CI lower bound .97) indicating excellent agreement. When the analysis was restricted to elderly residents and then repeated for stroke patients the respective ICCs were .92 (95% CI lower bound .85) and .98 (95% CI lower bound .96).

4.3.3. Intra-rater reliability

To assess intra-rater reliability, seven raters (5 physical therapists, 1 occupational therapist and 1 nurse), evaluated 24 stable subjects (18 elderly residents and 6 stroke

FIGURE 4.4

Paired Balance Scale ratings for subjects in the inter-rater reliability study (N=63)

Figure 4.4.1. Paired ratings of the elderly residents (N=28)

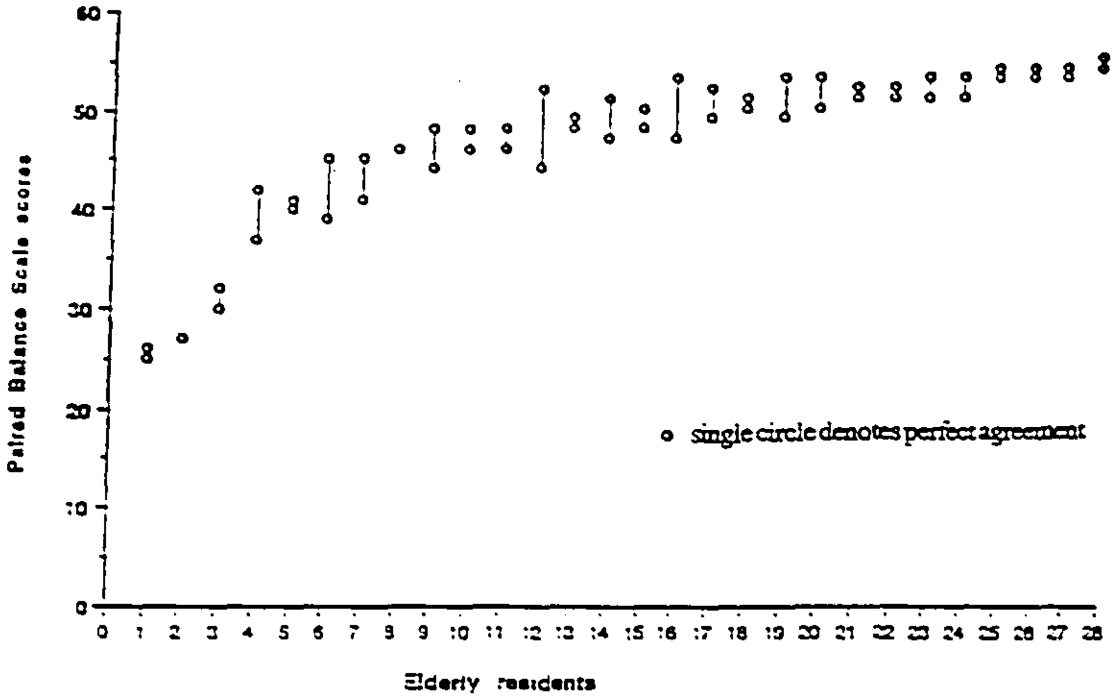
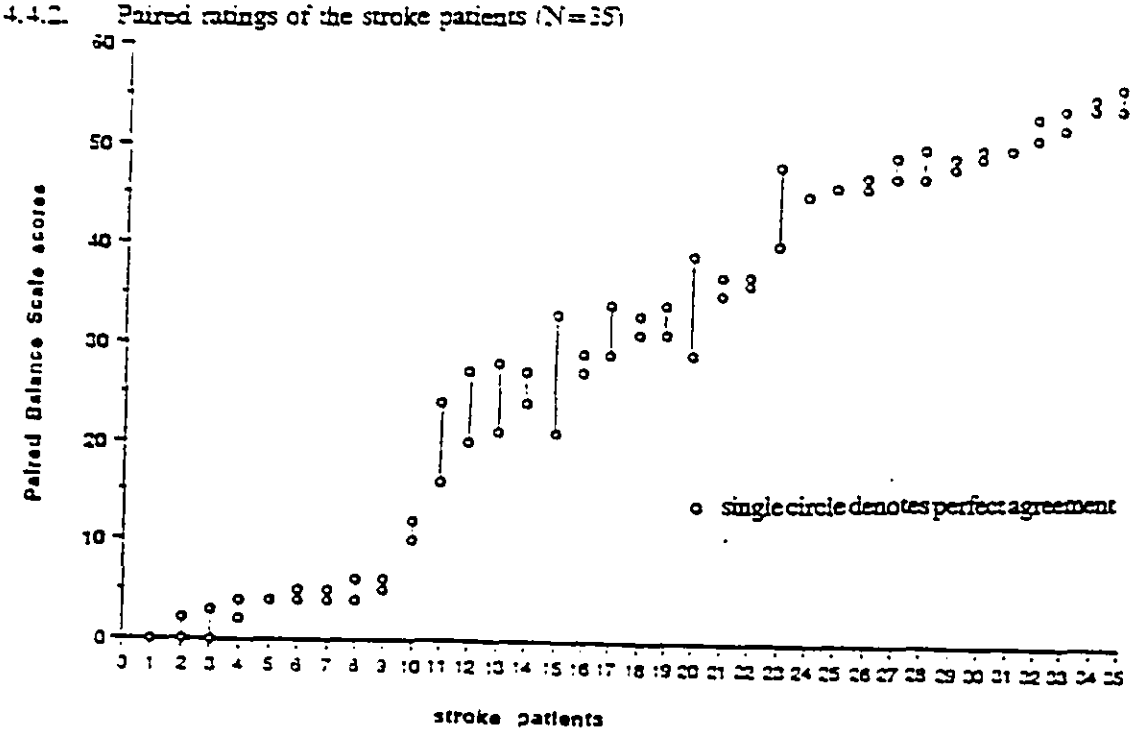


Figure 4.4.2. Paired ratings of the stroke patients (N=35)



	ICC	Lower bound of 95% CI
All subjects	.98	.97
Elderly residents	.92	.85
Stroke patients	.98	.96

patients) twice, one week apart. The range of Balance Scale scores was from 4-56 with an average of 46.0 (SD 11.0) over both ratings. The pairs of ratings are presented in Figure 4.5.

The ICC was used to assess the level of agreement using information from a two-way analysis of variance with subjects and time as factors. The ICC for all subjects was .97 (95% CI .93 -.99); whereas elderly residents showed an ICC of .91 (.80 -.96) and stroke patients an ICC of .99 (CI .94 -.999).

4.3.4 Internal consistency

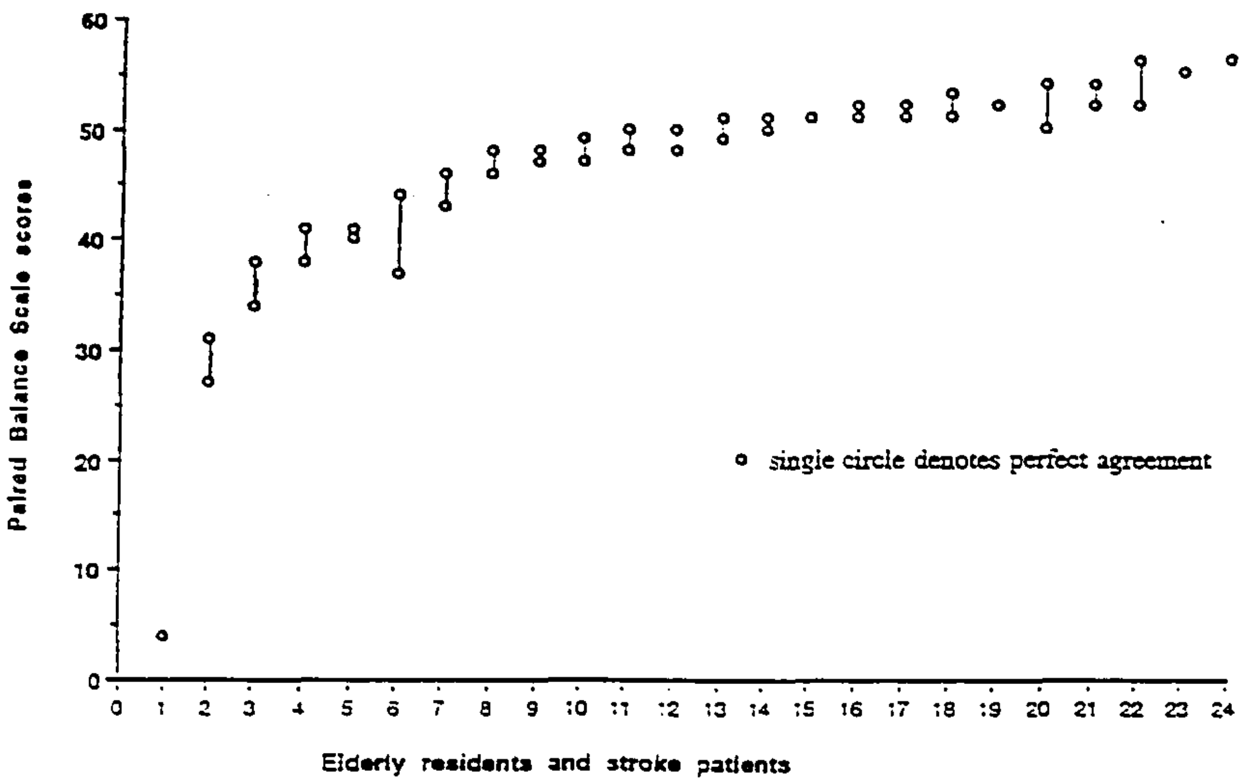
Internal consistency analyses were performed on the data from both stroke patients and elderly residents at each evaluation point. The correlation matrices including all items, except sitting unsupported, are displayed in Table 4.12. Because all elderly residents could sit unsupported, the relationship of this item to each of the others and the total score could not be tested.

Cronbach's Alpha and Standardized Item Alpha values were above .83 and .85, respectively, at each evaluation, suggesting that the scale is measuring one underlying concept and that all the items are contributing to the overall score. The item-to-total correlations for each of the four evaluations of elderly residents are shown in Appendix 4.6. Tandem standing and standing on one leg demonstrate item-to-total correlations below .4 on two occasions. This means that in the sample of elderly residents, a subject's performance on these items is not strongly related to scores for the remaining items. However, because these same items worked well for stroke patients and for the other two evaluation points, it was decided to retain them.

The internal consistency of the Balance Scale in the stroke population was even higher than in elderly residents. Cronbach's Alpha and Standardized Item Alpha values were consistent (.97 to .98) at each evaluation and all item-to-total correlations were above .62 (Appendix 4.7). The higher correlations in this population may be partially explained by the greater intersubject variability in performance. Uniformity of scores across subjects makes it difficult to obtain a high coefficient for a correlation between two variables.

FIGURE 4.5

Paired Balance Scale ratings of the elderly residents and stroke patients in the intra-rater reliability study (N=24)



	ICC	95% CI
All subjects	.97	.93 to .99
Elderly residents	.91	.80 to .96
Stroke patients	.99	.94 to .99

TABLE 4.12

Inter-item correlations of the Balance Scale at the initial evaluation

ELDERLY SUBJECTS

	SIT-STAND	STAND	STAND-SIT	TRANS	STEC	STFT	ARM REACH	PICK UP	TWIST TURN	TURN 360°	STEP STOOL	TANDEM
Standing	.36											
Stand to Sit	.69	.37										
Transfer	.68	.27	.64									
Stand Eyes Closed	.56	.54	.54	.47								
Stand Feet Together	.45	.19	.44	.36	.33							
Arm Reaching	.41	.44	.41	.28	.22	.17						
Object Pick Up	.36	.44	.28	.34	.23	.15	.28					
Twist Turn	.56	.32	.52	.49	.62	.37	.27	.42				
Turn 360°	.44	.20	.42	.40	.47	.10	.31	.49	.54			
Step on Stool	.36	.15	.39	.41	.28	.30	.35	.38	.37	.45		
Tandem Standing	.15	.10	.22	.20	.09	.25	.21	.28	.18	.26	.35	
One Leg Standing	.24	.14	.20	.22	.25	.14	.24	.24	.32	.37	.40	.23

ALPHA = .83
STANDARDIZED ITEM ALPHA = .87

average r	minimum	maximum	percent below .3
.34	.09	.69	41%

STROKE PATIENTS

	SIT-STAND	STAND	SIT	STAND-SIT	TRANS	STEC	STFT	ARM REACH	PICK UP	TWIST TURN	TURN 360°	STEP STOOL	TANDEM
Standing	.88												
Sit	.68	.73											
Stand to Sit	.88	.90	.66										
Transfer	.90	.89	.71	.83									
Stand Eyes Closed	.83	.92	.69	.86	.87								
Stand Feet Together	.82	.81	.58	.79	.81	.78							
Arm Reaching	.88	.87	.62	.87	.88	.81	.89						
Object Pick Up	.76	.82	.55	.83	.79	.77	.80	.79					
Twist Turn	.84	.88	.63	.89	.86	.85	.88	.92	.90				
Turn 360°	.74	.77	.50	.74	.80	.76	.81	.81	.87	.86			
Step on Stool	.63	.65	.40	.65	.67	.60	.72	.68	.77	.73	.78		
Tandem Standing	.65	.70	.46	.70	.74	.71	.80	.76	.77	.78	.77	.73	
One Leg Standing	.51	.60	.40	.62	.63	.60	.62	.69	.72	.74	.66	.70	75

ALPHA = .97
STANDARDIZED ITEM ALPHA = .98

average r	minimum	maximum	percent below .3
.75	.40	.92	0%

The mean scores for each item performed by the elderly residents remained relatively stable over four evaluations (Appendix 4.8). The hierarchy for item difficulty is similar to that found in the stroke sample. In both samples, the magnitude of the intra class correlation coefficients and Cronbach's alpha are considered to be high.

Summary for Reliability Study

The results showed excellent agreement for both inter and intra-rater reliability with intraclass correlation coefficients of .98 and .97, respectively. In addition, standardized Cronbach's alpha estimates were high in both elderly residents (.87) and stroke patients (.98), indicating strong internal consistency.

4.4 Chapter Summary

Tables 4.13 and 4.14 summarize the measurement properties of the Balance Scale as accrued from the preliminary study that developed the content of the scale (Berg et al. 1989), from the study that compared clinical and laboratory measures of balance (Berg et al. 1992) and from the present measurement study.

The present investigation examined the measurement properties of the Balance Scale in three inter-related studies. Evidence for the criterion validity of the scale is provided by the moderate association with clinical global ratings of balance, the relationship between Balance Scale scores and falling in the study period, and the ability of the Balance Scale to differentiate between groups on the basis of type of walking aid and location of the follow-up evaluation for stroke patients. The construct validity of the Balance Scale is supported by the covariation over the four evaluations of the Balance Scale scores with measures of functional status and motor performance in stroke patients. The similarity between the Balance Scale and the Barthel Index in detecting changes in the status of the stroke patients in the 12-week study period presents evidence of the instrument's responsiveness.

The third study assessed reliability, a prerequisite to both validity and greater responsiveness. Results demonstrated excellent agreement for both inter and intra-rater reliability in elderly residents and stroke patients. In addition, the internal consistency of the scale was strong, suggesting it measured one underlying dimension.

TABLE 4.13

Summary of validity assessments of the Balance Scale as tested by the author

STUDY and Subjects	VALIDITY: Content	Criterion	Construct
Berg et al. 1989 • 32 professionals • 38 geriatric subjects	• Formally developed using 3 panels of geriatric subjects and health professionals		
Berg et al. 1992 Cross-sectional comparisons of clinical and lab measures of 31 elderly subjects		• High correlation with Tinetti Balance Sub-scale • Moderate correlations with postural sway measures	
Present Measurement Study One-year longitudinal study of 113 elderly residents		• Moderate correlation with global clinical judgements of balance • Predictive of future falls by the elderly residents in the year of the study • Able to discriminate groups by: - type of walking aid used by the elderly residents, and	
Twelve-week longitudinal study of 60 stroke patients		- location of the follow-up evaluation of stroke patients	• Strong association with measures of functional status and motor performance in stroke patients over the 12-week period

TABLE 4.14

Summary of reliability and responsiveness estimates as tested by the author

STUDY and Subjects	RELIABILITY: Inter-rater	RELIABILITY: Intra-rater	RELIABILITY: Internal consistency	RESPONSIVENESS
<u>Berg et al. 1989</u> Raters: 10 professionals Subjects: 14 geriatric subjects on videotape	ICC = .98	ICC = .98	Cronbach's α = .96	
<u>Berg et al. 1992</u> Raters: 2 physical therapists Subjects: 10 geriatric subjects on videotape	ICC = .98			
Cross-sectional comparisons of 31 subjects performing clinical and laboratory tests of balance				• Balance Scale showed largest effect size in discriminating among subjects by their use of walking aids
<u>Present Measurement Study</u> Raters: 32 professionals and para-professionals Subjects: 31 elderly residents 36 stroke patients	ICC = .98	ICC = .97	Cronbach's α • elderly residents = .87 • stroke patients = .98	
Subjects for assessing responsiveness were 60 stroke patients who completed study				• Able to detect changes of same magnitude as Barthel Index

Chapter 5 DISCUSSION

5.0. Introduction

Three inter-related studies contributed information on how the Balance Scale performed in real life situations. The first two examined evidence of the validity of the instrument in terms of both criterion-related and construct validation strategies. The third study assessed three aspects of reliability: inter-rater, intra-rater and internal consistency.

This chapter summarizes the results and discusses their implications relative to the essential measurement properties of reliability, validity and responsiveness. The accrued information on the Balance Scale is also compared to what is known of the properties of other balance measures. Finally, the chapter discusses the limitations of the thesis work and suggests future directions for research and clinical practice.

5.1. Reliability

The discussion of reliability examines the evidence of inter and intra-rater reproducibility and internal consistency of the Balance Scale separately. Each section discusses the relevance of the findings and considers the minor differences observed in the results of two samples, elderly residents and stroke patients. Lastly, the generalizability of the reliability results is discussed.

Reliability is discussed first because it is a prerequisite to both validity and responsiveness. Validity coefficients cannot be interpreted without an estimate of the magnitude of the measurement error. Consequently, inadequate reliability generally precludes further testing or requires a revision of the instrument. In addition, without excellent reliability an instrument is unlikely to be responsive to small yet clinically important changes in status. Therefore reliability coefficients above .94 or .98 are recommended for instruments used to make decisions about individuals over two or more evaluations (Helmstadter 1964; Nunnally 1978).

5.1.1. Inter and intra rater reliability

The results showed excellent agreement in the paired ratings of the Balance Scale scores with intra class correlation coefficients for inter and intra-rater reliability of .98

and .97, respectively. The inter-rater assessment compared 63 paired ratings by 32 different raters who were nurses, senior matrons, physical or occupational therapists. Subjects included 35 stroke patients and 28 elderly residents. In addition, 24 subjects were evaluated at two points in time by seven different raters to assess intra-rater reliability. Each rating was made independently at time intervals that varied according to the stability of the status of the subjects. For example, the paired ratings for evaluating inter-rater agreement in stroke patients were made within 24 hours of each other; whereas, paired ratings of the elderly residents were made within one week.

Various sources of error can influence the reproducibility of scores. One of the first aspects examined is whether the scale itself may be poorly constructed and ambiguous, allowing subjectivity in scoring by the raters. This source of error can be addressed through fairly controlled testing conditions or by having raters evaluate videotaped performances where no actual changes can occur in the administration of the test or in the patient's performance.

In real life when an instrument is administered independently on two or more occasions, raters, patients and environmental conditions may be potential sources of error. Raters may differ in how closely they follow the written instructions. Caregivers may give higher marks than independent evaluators because they know the patient did the task well on a previous occasion. Patients may not perform the same way because of fear, fatigue, cognitive impairment or lack of motivation. A subject's performance may also differ depending on the evaluator. For example, they may perform better and feel more secure with someone they know as opposed to an independent evaluator. In addition, noise, visual distractions, unsuitable furnishings and other environmental conditions may contribute to measurement error in a performance-based instrument. The estimates of reliability from this study are more impressive because they include errors from a variety of sources that occur in real life.

The reliability of the Balance Scale is summarized by the intra class correlation coefficient. Measurement theory states that observed scores contain both real differences between subjects and random variation. Reliability is the proportion of the observed variance that is attributable to the true score differences between subjects. The intra class correlation coefficient measures this relative magnitude of inter-subject differences to the total variation in scores using estimates of variance derived from an analysis of variance. The smaller the error variance, the more the denominator will approximate the numerator and the closer the ratio will be to 1. Given equal error variance, the intra class correlation will also be larger when there is greater variation between subjects.

It is also helpful to verify the results with plots of the paired ratings, examination of the size of the mean square error, and the range of ability of the subjects. Using this verification approach, the high reliability coefficients appear justified. The plots of the paired ratings indicate good agreement on average. In addition, the magnitude of the error variance was small in both inter-rater and intra-rater analyses.

There were, however, minor differences in the estimates of reliability between the analyses of elderly residents and stroke patients. The estimates of inter-rater reliability of stroke patients were higher with an ICC of .98 (95% CI lower bound .96) relative to an ICC of .92 (95% CI lower bound .85) for the elderly residents. Similarly, intra-rater reliability estimates were an ICC of .99 (95% CI .94 - .99) for stroke patients and .91 (95% CI .80 - .96) for elderly residents. The variation may be explained by the difference in the range of ability in the two groups. Although the impact on the results in this study was minor because the range in the elderly was sufficient to be representative of this group, it illustrates a drawback of the ICC as a statistic for assessing reliability.

Close examination of the paired ratings and the magnitude of the error variance does not support greater reliability in the ratings of stroke patients. There were more numerous and larger discrepancies in the paired ratings of stroke patients, and the mean square error estimates were larger than the similar analyses of elderly residents. In each instance, however, elderly residents showed a narrower range of Balance Scale scores, indicating less variation between the subjects, a factor that decreases the numerator and the overall intra class correlation coefficient.

Another reason for higher estimates in the stroke population may be the multiple zeroes in the scores of low level patients. A rating of zero leaves less opportunity for error in the retest situation. Low functioning subjects are awarded the same scores with repeated tests because they clearly fail or cannot try difficult tasks. In contrast, elderly residents could attempt all tasks.

The plots of the paired ratings in stroke patients suggest that greater discrepancies in scores occurred in the middle ranges of ability. This finding may reflect an inherent inconsistency in the performances of some patients during the early recovery period. Stroke patients must adjust to their deficits and relearn basic motor skills, a process that is associated with variability in performance. This variability may be more noticeable for patients who can attempt all items but vary in the degree to which they can meet the scoring criteria. A better estimate of their ability would be the average of two or more tests.

Instability in performance may also explain why it was difficult to identify stroke patients in the middle range of ability who were sufficiently stable to be retested in a week. Consequently, subjects in the intra-rater study do not represent the full range of scores on the Balance Scale. Further testing should make a greater effort to include patients in the middle to lower end of the range of the Balance Scale.

5.1.2. Internal consistency

The Balance Scale shows high internal consistency with Cronbach's alpha averaging .87 for the elderly residents and and .98 for the stroke patients. Assessment of internal consistency was made separately for the stroke patients and elderly residents at each evaluation using all available subjects. This strong internal consistency facilitates the interpretation of scores but it is not essential for good measurement. It is possible to have a multi-dimensional scale that is reproducible and valid for different purposes. The primary advantage of having multiple homogeneous items in the Balance Scale is that it provides a basis for a more consistent estimate of the ability of individuals to balance.

Cronbach's alpha was developed as a reliability estimate for multi-item tests or indices. The assumption underlying testing of reliability based on a single administration of the test is that each item may be considered a single measure of a common underlying characteristic, and the sum of these related items should be more reliable than any item individually (Bravo and Potvin 1991). Because the items should be correlated with each other and with the total score to capture the concept of interest, this form of reliability is called internal consistency.

The greater the number of items and the more the items covary relative to the sum of their variance, the higher will be the Cronbach's alpha. Cronbach's alpha is considered a lower bound of reliability unless the items in the scale are parallel, having equal means and variances and equal correlations with a third variable (Bravo and Potvin 1991). Given the stringent definition of parallel items, it is expected that the correlation between the same test administered at two separate occasions in the absence of true change should be higher than among multiple items in a scale.

The results of this study are consistent with that expectation; the intra class correlation coefficients are higher than the Cronbach's alpha's in each situation. However for each group, elderly residents and stroke patients, the estimates of internal consistency are within the 95% confidence interval of the inter and intra-rater reliability estimates. The overlapping of the confidence interval is indicative of the shared assumptions of the two types of reliability coefficients, both being based on the same measurement model and the same definition of reliability (Bravo and Potvin 1991).

The variation in magnitude of the inter-item correlations within each sub-study illustrates the importance of testing an instrument in diverse populations. At entry to the study, the inter-item correlations for stroke patients ranged from .40 to .92 whereas those in the elderly residents ranged from .09 to .69. Although the data for the elderly residents showed much lower inter-item correlations, the Cronbach's alpha remained high (.87), indicating that each item was contributing additional information to the measurement of the concept and that the group of items was internally consistent. Based solely on data from the stroke patients, certain items with inter-item correlations greater than .90 could be deleted from the scale as they offer redundant information.

The greater internal consistency estimates in the study of stroke patients may also be related to the wider range of ability in this group. The scoring of most items used each of the five response categories. In contrast, the frequency distribution and the mean scores for each item indicated that most residents scored in the higher response categories on the majority of items. This lack of range within each item may have masked the true relationship between items and thus, artificially lowered the Cronbach's alpha.

When verifying the composition of a scale, it is also useful to examine how well each item correlates with the sum of the scores of the remaining items in the scale. In the data pertaining to the stroke patients, all the item-to-total correlations were above .63. The item-to-total correlations were also satisfactory in the study of the elderly residents except for two items which had estimates hovering around .40, the criterion recommended for retaining items. The items required subjects to stand on one leg and stand with one leg in front of the other. Judging by the mean scores for each item, these were the most difficult tasks to perform. The lack of a strong association with the total of the remaining items may relate to the sharp increase in the degree of difficulty, when compared to the other tasks, rather than their inappropriateness as items of balance. Both single leg and tandem stance are widely used as tests of balance in research and clinical practice (Bohannon 1984; Briggs et al. 1989; Goldie et al 1989; Heitman et al. 1989; Lichtenstein et al 1989).

5.1.3. Generalizability of the reliability results

The findings should be applicable to most clinical situations. The generalizability of the results is strengthened by the varied clinical characteristics of the subjects, the diversity and lack of training of the raters, and the lack of control of the test conditions.

The majority of caregivers participating in the study were physical therapists, 66% and 71% of the raters in the inter and intra-rater reliability studies, respectively. They reflected diverse levels of experience. Some were student therapists and others had been working more than 30 years. The other professionals were either nurses or occupational therapists. In addition, at the home for the elderly, two paraprofessionals called senior matrons participated in the study. The raters received no formal training in the administration of the Balance Scale but they were asked to read through all the items and ask questions as necessary. Although the random pairing of the raters did not permit a separate estimation of rater influences or variance due to a particular profession, this variation is included as part of the error mean square.

Testing of the Balance Scale in conditions that simulated clinical reality supports its use in clinical practice and research. Paired tests were made at different times of the day, often in two separate locations with different furnishings and noise levels or other distractions. Despite the possible sources of variation, the Balance Scale demonstrated high reliability. Discrepancies of possible concern to clinicians were few in number but could be addressed in the future by repeating the test and averaging the results whenever an inconsistent performance is suspected.

5.2. Evidence of the validity of the Balance Scale

Four measures, external to the Balance Scale, were used to examine its criterion validity. The scores on the Balance Scale were compared to global judgments of balance, the occurrence of falls, use of mobility aids, and the location of the follow-up evaluation of the stroke patients. Each criterion is related to the concept of balance and together, the associations accrue evidence for the validity of the scale.

In addition, the evidence for the construct validity and responsiveness of the Balance Scale is discussed in the context of the performance of the Balance Scale in monitoring the status of stroke patients. The theoretical basis underlying the use of construct validity was the expected covariation among measures of motor performance, functional status and balance in patients following an acute stroke. The anticipated changes in the status of the patients in the 12 week period also provided the opportunity to examine the responsiveness of the Balance Scale.

The generalizability of the validity testing is discussed at the end of the section.

5.2.1. Relationship of Balance Scale scores to global judgments of caregivers

Balance Scale scores were compared at baseline, 3, 6 and 9 months to concurrent global ratings of balance given by caregivers. The judgments of the caregivers were made within 48 hours of the independent evaluator's rating of the Balance Scale. The correlations between the two evaluation methods ranged from .50 to .62. The moderately high correlation with global ratings showed that the scores on the Balance Scale were in agreement with the judgments of clinicians, the potential users of this scale. The use of global ratings is common in clinical practice, although the need for more objective quantitative methods to assess the effectiveness of interventions and monitor the status of patients is becoming more evident.

The importance of showing this association over-shadowed concerns about the measurement properties of this criterion. Little is known of the reliability and validity of global ratings. They lack precision, having only three options, good, fair and poor, making it difficult to identify small differences between subjects and improvements in ability over time. It is doubtful that they can be easily compared between professionals or institutions with different patient populations because the raters can use different reference criteria to assign categories. This lack of information on the magnitude of measurement error makes it difficult to interpret validity coefficients because unreliability attenuates the correlation. Therefore, the moderate association between the Balance Scale scores and the global ratings is likely an underestimate of the true relationship between the variables.

The comparison between Balance Scale scores and global ratings was also limited by the range of ability of the subjects. At baseline no subject scored below 23 on the Balance Scale and only 10 (8.8%) of the residents were rated as having poor balance. The small numbers within the poor category at any given time also contributed to the fluctuations in the mean scores of this group and the wide variability at each evaluation. The present study is, however, reflective of evaluations made in the community on subjects who are reasonably independent in the basic activities of daily living and free of acute illnesses.

The moderately high correlations found between Balance Scale scores and caregivers' global judgments are consistent with the process of content development in which three panels of clinicians and patients participated (Berg et al. 1989). Items were chosen for their relevance to balance and the subsequent evidence of reliability. Global judgments are likely based on the observations of the performance of similar items. Every day, the caregivers saw the residents going to meals, manoeuvring around the tables in the dining area, and standing up and sitting down for meals.

The evidence in favour of a positive association is strengthened by the consistency of the correlation coefficients at subsequent evaluations. It is also supported by the gradient of mean Balance Scale scores in the good, fair and poor categories that remained greater than expected by chance when the Barthel Index scores and age of the subjects were used as covariates. Nonetheless, the association with global ratings of balance cannot stand alone as evidence of the validity of the Balance Scale. Comparisons must be made with other measures of balance. Further information on what is known relative to other tests is discussed in a later section of the chapter.

5.2.2. Relationship of the Balance Scale to the occurrence of falls

Falling was selected as an external criterion because it is an objective sign of loss of balance and is of itself an area of concern in geriatrics. However, it was recognized that falling and balance do not have a straightforward relationship. Falls are a multifactorial problem and should not be predictable on the basis of a balance score alone. Nevertheless, there should be an association whereby knowledge of Balance Scale scores would improve the prediction of falls.

The results demonstrated that poorer Balance Scale scores were associated with falling. There was a weak to moderate (rho=-.36) negative correlation between Balance Scale scores at baseline and fall status as defined by no falls, a single fall and two or more falls in the year of follow-up. The risk of falling was also increased in individuals with Balance Scale scores under 45 compared to those with scores of 45 and greater (RR one fall 1.5; RR two or more falls 2.7). In addition, the Balance Scale scores were an important contribution to predicting falls in the logistic regression. The three methods used in the analysis are complementary and support the predictive validity of the Balance Scale in terms of forecasting falls.

The present study, however, was not primarily designed to identify the risk factors of falling. Repeated tests of balance were performed at 3-month intervals, too long a period in most cases to assess the effect of changes occurring immediately prior to a fall. Nonetheless, the repetitions did permit a verification of the stability of the Balance Scale scores in most subjects over the year of follow-up.

The report on falls in the present study is consistent with statistics reported in the literature. The proportion of individuals falling in the year was 47%, higher than the 32% reported in a community based study (Tinetti et al. 1988), but close to the 45% figure found in a similar residential care facility (Gryfe et al. 1977). The proportion of falls that resulted in a fracture was 6%, consistent with previous findings (Gryfe et al.1977; Tinetti et al. 1988). Furthermore, balance, when measured by Functional Reach (Duncan et al. 1992) and when assessed as a performance-based scale (Tinetti et al. 1988), has previously been reported as a risk factor for falling.

The presence of visual deficits and history of falls were also identified as risk factors for more than one fall in the year of study. Each has been associated with risk of fractures or falls in previous research (Felson et al. 1989; Mayo et al 1989; Morse et al. 1987). The history of falls was also determined by self-report and considered only relative to the three months prior to enrolment in the study. Any falls that were known to have occurred outside of that period were not counted.

Several other variables have been reported to be risk factors for falls in the elderly. They include medications (Grisso et al. 1991; Ray et al. 1987; Tinetti et al. 1988), cognitive impairment (Morse et al. 1987; Tinetti et al. 1988), age (Campbell et al. 1990; Vlahov et al. 1990), and specific diagnosis (Grisso et al. 1991; Lipshitz et al. 1991; Mayo et al. 1989). Each of these variables was measured in the present study but did not contribute to the prediction of falls.

5.2.3. Relationship to other external indicators of balance

The two most commonly asked questions by clinicians are whether Balance Scale scores can help them reach decisions about the most suitable mobility aid and the most timely date for a safe discharge home. Future research will determine whether the Balance Scale can be effectively used to make clinical decisions. The cross-sectional portion of the analysis, however, focused on the ability of the Scale to discriminate among groups expected to differ on these two important external criteria: use of mobility aids and location of the evaluation following an acute stroke. The latter is essentially a measure of whether the subjects had already been discharged home, to a rehabilitation centre or remained confined to the general hospital to which they were admitted for their stroke.

The results of both comparisons showed that the mean scores showed a linear trend in the expected direction, indicating that the Scale could differentiate groups according to their use of mobility aids and the location of their follow-up evaluation. Of the two, discriminating between the types of mobility aid was a greater challenge to the Balance Scale because the difference in the mean Balance Scale scores between the groups using no aids (49.5), canes outdoors only (48.3), canes (45.3) and walkers (33.1) was smaller than in the comparison with the different locations of the evaluation. The elderly residents were all medically stable and independently mobile at entry to the study. Their use of mobility aids was likely voluntary, based on recommendations of staff or on their own concerns about safety. In addition, the association with the use of a mobility aid could be obscured if individuals did not use prescribed walking aids or had ones that were inappropriate for their level of ability. No efforts were made to determine why they used the aids or to verify reported use. Regardless, the mean scores showed the expected linear trend and the difference in mean scores remained after Barthel Index mobility scores, age and orthopeadic diagnosis were taken into account in the analysis. Discriminating among the groups with small yet clinically significant differences required greater precision on the part of the Balance Scale, a quality that is advantageous for an outcome measure.

At each follow-up evaluation there were marked differences in mean scores based on the location of the assessment. The patients discharged home had mean scores of 40 or above each time. Those in a rehabilitation centre had mean scores that ranged from 27.3 to 32.9. Hospitalized patients at 4 weeks had a mean of 19.5, but, by 12 weeks, those still in general hospital had a mean score of only 8.1. Earlier in the recovery period other factors such as living arrangements, associated medical conditions, and occupancy levels at rehabilitation hospitals may have slowed their release. By 12 weeks, only patients identified as poor rehabilitation candidates are left in general hospital.

Whether or not the Balance Scale total scores prove useful in making predictions about safe mobility and safe return home for an individual patient, the component items are those that currently help clinicians to make their decisions. It is important to know how well a patient can stand up, turn around, and transfer. An advantage of the Balance Scale is that it is composed of clinically relevant items.

5.2.4. Performance of the Balance Scale in monitoring the status of stroke patients

Acute stroke patients with motor and functional deficits are generally expected to improve and thus, provide an ideal population in which to assess the performance of the Balance Scale. At each of the four evaluations, balance scores were highly correlated with measures of functional status and motor performance. Specifically, correlations between Balance Scale and total Barthel Index scores varied from .87 to .93 over the four evaluations, whereas correlations between the Balance Scale scores and total Fugl-Meyer scores were .70 to .82. It is anticipated that the association with motor performance would be lower because patients can have movement without having sufficient postural control to use it functionally. This fact is illustrated by the higher mean scores for all stroke patients on the Arm and Leg sub-scales of the Fugl-Meyer (47.3) relative to either the Barthel Index score (33.8) or the Balance Scale score (27.7) when each measure is scored out of 100.

The relationship between the measures was strengthened by the way the Balance Scale scores correlated with the sub-scales of the Barthel Index and Fugl-Meyer Scale. The strongest associations were found between the Balance Scale and the Barthel Index mobility section (.94) and the Fugl-Meyer Balance sub-scale (.94) at the 12-week evaluation. The lowest was with the Fugl-Meyer Arm sub-scale (.62) at entry to the study. The relationship may be weakest initially because subjects have not yet learned to adapt their postural control to the stroke-related deficits. In addition, the component tasks of the arm sub-section are performed while sitting, a position with a broad base of support that requires less postural control. Nonetheless, there should be a relationship between Balance Scale scores and the Fugl-Meyer arm sub-scale because postural adjustments are needed for any voluntary movement.

The correlations between the Balance Scale and the lower extremity portion of the Fugl-Meyer Scale at each evaluation were intermediate to the upper extremity sub-scale and the balance sub-scale. This finding is consistent with the expected impact of lower extremity weakness and diminished coordination on the ability to balance. However, given that none of the lower extremity tests of the Fugl-Meyer require that the subject stand unsupported, balance dysfunction does not unduly influence the scores.

The pattern of covariation of the scores of balance, functional status and motor performance supports the longitudinal construct validity of the Balance Scale and provides initial evidence of its suitability in monitoring the clinical status of the patients over time. Further evidence is provided by the direct comparison of the magnitude of change shown by the Balance Scale and Barthel Index scores over the course of the study. This analysis involved computing the difference between the Barthel Index and Balance Scale scores for each subject, and testing whether this difference score remained constant over the four evaluations. The results indicate that the Balance Scale is at least as sensitive in detecting changes over time in this sample as the Barthel Index. In contrast, although able to show changes over time, neither the combined Arm and Leg sub-scales of the Fugl-Meyer or the Fugl-Meyer Balance Sub-scale were able to detect the same magnitude of change in subjects in the 12-week study period.

In addition, the greater responsiveness of the Barthel Index relative to the Fugl-Meyer Motor Performance Scale is consistent with previous studies. Wood-Dauphinee and associates (1990) compared a neurological assessment measure, a stroke severity scale, the Fugl-Meyer Scale and the Barthel Index in their ability to detect a treatment effect in acute stroke patients. The Barthel Index showed the largest effect size, an indication of its superior responsiveness.

In acute stroke patients, the Barthel Index is a good criterion measure of responsiveness. The Index is widely used and recently was recommended as the best measure of activities of daily living (Wade 1992). Furthermore, an examination of its content shows how pertinent the items of the Barthel Index are to an acute, dependent group of subjects such as the participants of this study at baseline. To achieve a degree of independence at the end of 12 weeks, the subjects had to show improvement in tasks that are critical to independent function and are important markers for the improvement of stroke patients in the early recovery period.

Although not previously tested, the responsiveness of the Balance Scale was considered during the phases of content development. Professionals were asked to suggest response categories that were clinically relevant. The gradations in responses, included in the final version of the instrument, were based on the independence-dependence continuum. For example, subjects are awarded full marks if they are able to perform the task independently within a certain time allottment. Progressively fewer points are awarded if time constraints are not met or if greater supervision, cueing, or assistance is required. The improvements seen in the mean scores of each item in subsequent evaluations of the stroke patients suggest that each item is able to contribute to monitoring the status of the patients.

It is also encouraging that the Balance Scale does not show clustering at the top of the scale. It was able to detect changes of the same magnitude as the Barthel Index but the Balance Scale scores were proportionally lower at each evaluation, suggesting there is room for continued improvement in scores. Further studies are necessary to determine its ability to monitor change in other groups and to discriminate differences in change scores between a treatment and control group in a clinical trial.

5.2.5. Generalizability of the results of the validation studies

The use of different subjects and strategies to assess the validity of the Balance Scale has provided a substantial amount of information on its performance. Balance Scale scores are compatible with clinical judgements. Additionally, they are associated with an elevated risk of falling, use of mobility aids and discharge location in the recovery period of stroke patients. Balance Scale scores also show a strong relationship to two related constructs, motor performance and functional status. The covariation of the three measures over the study period provides evidence of construct validity. Lastly, the Balance Scale is able to monitor changes in the status of stroke patients to the same degree as the Barthel Index.

Overall, the subjects in the validation studies representated a wide range of ability, a variety of medical conditions and an age spread from 44 to 97. Elderly subjects appeared representative of their institution but it is difficult to judge the extent to which they are similar to the elderly at large. Their fall statistics are compatible with previous studies (Gryfe et al. 1987; Tinetti et al. 1988). The characteristics of the stroke patients suggest that they are similar to patients with motor and functional deficits who are referred to rehabilitation. Moreover, there were few inclusion or exclusion criteria to bias the samples.

The Balance Scale has performed to expectation in each aspect of its testing and is ready for use in clinical practice and research. While not all questions concerning its measurement properties for different situations have been answered, the consistency of the findings support its use and its generalizability.

5.3. Limitations of the study

The assessment of reliability showed excellent agreement using patients with varied characteristics and raters with diverse professional backgrounds. The professionals received little training in the administration of the instrument, a fact that strengthened the generalizability of the findings. However, most raters were physical or occupational therapists. The small numbers who were nurses or para-professionals and the study design did not permit an assessment of the variation specific to the profession. In addition, researchers have inquired about using the Balance Scale in community-based research where it could be administered by non-professionals. We do not yet know how much training may be required for non-professionals.

As previously mentioned, there may be greater inconsistencies in the performance of patients in the middle ranges of ability. Further investigation can examine whether this is the case and, if necessary, assess the number of tests that must be averaged to give a reliable result (Fleiss 1986).

However, the primary limitations to the present study are fundamental to the on-going process of instrument validation in the absence of a gold standard. Assessment of validity involves judgments about which criteria are relevant to the concept and what magnitude of correlation is required to support the validity of the instrument.

There are problems associated with each of the criteria employed in the present study. Global ratings are subjective and have no documentation of their reliability or validity. Nonetheless, they are important for showing that clinical judgments are consistent with scores on the Balance Scale. Falling is a multi-factorial problem that is unquestionably related to balance but the relationship is not straightforward. Individuals may compensate for their balance deficits with the use of mobility aids or by restricting their opportunity to fall. In addition, use of mobility aids and location of the follow-up evaluations are indicators of different levels of ability but they are also influenced by other factors. Similarly, examining the evidence of construct validity requires a judgment as to the nature of the expected relationship between balance and functional status and motor performance in stroke patients.

The suitability of the Balance Scale as an outcome measure requires that it demonstrate an ability to detect change in the patient's status with repeated administrations, discriminate small but clinically meaningful changes, and that it be able to discern a beneficial effect of treatment in a clinical trial. To date the Balance Scale has been able to discriminate between groups in the use of mobility aids and monitor changes in stroke patients to the same degree as the Barthel Index. Further information is needed to determine its ability to monitor changes in other subjects or to detect a treatment effect in a clinical trial.

While none of the strategies for examining validity can stand alone, together they are convincing of the validity of the Balance Scale in describing and discriminating subjects of varying degrees of balance dysfunction, assessing individuals for appropriate interventions and monitoring their status over time.

5.4. Summary of the comparisons with other measures of balance

The Balance Scale was developed for use with individuals who have a degree of impairment. The results to date have confirmed that there is sufficient range in the Scale to assess most patients undergoing rehabilitation. For example, acute stroke patients scored along the full available range (0-56). Hence, the scale could evaluate subjects from the earliest days of recovery to the time when they are living independently in the community. In contrast to other tests (Duncan et al. 1990; Horak and Shumway-Cook 1986; Maki et al. 1987; Wolfson et al. 1986), subjects do not have to be able to stand independently to score above zero. In addition, the Balance Scale showed differences in scores among elderly residents who were independent in functional status and had Barthel Index scores of 100.

Tables 4.15 and 4.16 compare the test properties of clinical measures of balance. To assess their validity, clinical measures have been compared to laboratory tests of sway (Berg et al. 1992; Dettman et al. 1987; Duncan et al. 1990; Lichtenstein et al. 1990). Laboratory measures, however, are not included in Table 4.13 because of their restricted appplicability. Postural sway tests generally require expensive equipment and trained personnel, making them impractical for use in clinical practice and many research projects. The Postural Stress Test (Wolfson et al. 1987) also requires equipment, lacks portability and, like the tests of postural sway, has questionable functional relevance. The Fugl-Meyer Balance sub-scale was omitted from the table because it has only been tested in the stroke population.

Although included in Tables 4.15 and 4.16, the Balance Coding (Gabell and Simons 1982) and the CTSIT (Horak 1987; Shumway-Cook 1986) have limited information of their measurement properties and neither is summed to provide a total score. The Balance Coding was the first performance-based scale but there are no published reports of its use in other settings. The CTSIT is recommended as a clinical assessment of the sensory constraints contributing to postural instability (Shumway-Cook and McCollum 1991).

More is known of the measurement properties of the Balance Scale, the Tinetti Balance Sub-scale (Tinetti 1986) and the measure of Functional Reach (Duncan et al. 1990). All three measures have demonstrated concurrent criterion validity relative to laboratory measures and predictive validity as regards falls in the elderly. Direct comparison between the Balance Scale and the Tinetti Balance Sub-scale showed a strong correlation (.91), indicating that at any one point in time they are providing similar information (Berg et al. 1992). Scores on the Balance Scale and the measure of Functional Reach have not yet been compared, but this item is included in the content of the Balance Scale.

When compared to the Balance Scale, the Tinetti Balance Sub-scale has weak

TABLE 4.15

Scoring methods and reliability assessments of clinical measures of balance

	Scoring (range of scores)	Method of Assessment	Subjects	RELIABILITY Inter-rater	RELIABILITY Intra-rater	RELIABILITY Internal Consistency
Balance Scale • Berg et al. 1989 • Berg et al. 1992 • Present Study	14 items (0-56)	Response categories based on time, distance and level of supervision in performing task	• Geriatric subjects (in-patients to community dwelling) • Stroke patients	Yes	Yes	Yes
Functional Reach • Duncan et al. 1990 • Duncan et al. 1992	Single item	Average of 3 tests of reaching while standing. Yardstick attached to wall at shoulder height.	• Community-dwelling elderly	Yes	Yes	Yes
Tinetti Balance Sub-Scale of Mobility Score • Tinetti 1986 • Tinetti et al. 1986 • Tinetti et al. 1988 • Lichtenstein et al. 1990	13 items (0-24)	Responsive categories based on subjective observations	• Residents of intermediate care facilities • Community-dwelling elderly	Yes (sparse)	No	No
CTSIT • Shumway-Cook & Horak 1986 • Horak 1987	No total score	Observations of postural sway while standing 30 sec in six sensory conditions	• Not actually tested but recommended for rehabilitation patients	No	No	No
Balance Coding Gabell and Simms 1982	Alphanumeric profile (maximum 6 ABX)	6 hierarchical standing tasks and 3 dichotomously graded tasks	• Geriatric subjects (in-patients to community-dwelling)	No	No	No

TABLE 4.16

Criteria used to assess validity of clinical measures of balance

	VALIDITY				RESPONSIVENESS
	Content	Criterion		Construct	
		Concurrent	Predictive		
<u>Balance Scale</u> • Berg et al. 1989 • Berg et al. 1992 • Present Study	Yes	• Global clinical judgements • Laboratory measures of sway • Tinetti Balance Sub-Scale • Use of mobility aids • Location of follow-up evaluation of stroke patients	Falls	• Covariation with functional status and motor performance in stroke patients	• Changes compared to Barthel Index • Effect size when discriminating subjects by use of walking aids
<u>Functional Reach</u> • Duncan et al. 1990 • Duncan et al. 1992 • Weiner et al. 1991 (abstract)	N/A	Laboratory measure of centre of pressure excursion	Falls	No	• Changes in male veterans following rehabilitation
<u>Tinetti Balance Sub-Scale of Mobility Score</u> • Tinetti 1986 • Tinetti et al. 1986 • Tinetti et al. 1988 • Lichtenstein et al. 1990	No	• Laboratory measures of postural sway • Balance Scale	Falls	No	No
<u>CTSIT</u> • Shumway-Cook & Horak 1986 • Horak 1987	No	No	No	No	No
<u>Balance Coding</u> Gabell & Simms 1982	No	• Differentiates between in-patients and out-patients	No	No	No

evidence of inter-rater reliability and lacks information on intra-rater reliability and internal consistency. It was primarily developed to predict falls and thus has not been encouraged as a descriptive measure or an outcome measure. Moreover, it was shown to have a smaller effect size relative to the Balance Scale when discriminating between groups of elderly individuals using no walking aids, a cane or walkers (Berg et al. 1992). A larger effect size is an indicator of greater power in detecting a true difference. This greater efficiency of the Balance Scale, as regards sample size, supports its potential as an outcome measure.

The lesser precision of the Tinetti Balance Sub-scale may be explained by its more limited scoring categories. For example, the responses in the Tinetti Balance Sub-scale consist of two or three choices that are descriptive in nature, such as taking discontinuous steps when turning in a circle. In contrast, the same item in the Balance Scale has five response categories, scored relative to time standards and level of supervison required for safety.

There are indications that the measure of Functional Reach is able to detect change in male veterans undergoing rehabilitation (Weiner et al. 1991). As in the case of the Balance Scale, the degree to which it can discriminate degrees of change between a treatment group and comparable control group has yet to be shown. One potential problem is that increasing reach by a couple of inches has little direct functional relevance. While it is unlikely that this test is sufficiently meaningful to be the principle outcome measure of a study, it may be very useful as a screening measure for individuals at high risk of falls.

Functional reach has recently been assessed for its ability to predict recurrent falls (Duncan et al. 1992). Relative to subjects who could reach forward 10 or more inches, individuals unable to stand independently had an adjusted odds ratio of 8.07 (2.8-23.7) for having recurrent falls in a six month follow-up period. Comparisons with the present study are limited by slight differences in the ordinal scoring of forward reach relative to the gradations of forward reach in the Balance Scale. Duncan and associates (1992) also included subjects with much lower levels of ability. Future examinations can more carefully compare the single item versus the total Balance Scale score.

A single item test, however, is unlikely to adequately assess a concept (Nunnally 1978). The advantages of multi-item tests are that they offer a more comprehensive assessment of the concept, improve reliability and allow for greater distinction between subjects (Bravo and Potvin 1992; Nunnally 1978).

On occasion, it is tempting to use only selected items from scales. This practice seems to sometimes occur if there is a specific purpose to the measurement. For example, Tinetti and associates (1988) chose four of the 13 items in the Balance Sub-scale as the best predictors of falling in community dwelling elderly. There are several potential problems with this approach. The selected items may not be the most appropriate for all groups. There may be a marked drop in the measurement properties. Most importantly, there is a loss in ability to describe and communicate the characteristics of the subjects to other professionals and researchers. The Balance Scale has been shown to be internally consistent and to have good measurement properties for the total score. Its use by institutions across Canada and other countries helps facilitate communication.

5.5. Future directions for research and practice

The Balance Scale was developed with input from health care professionals and geriatric subjects. The measure appears acceptable to both groups. Patients are willing to perform the items and have sometimes continued trying to improve their performance on an item long after the score has been recorded. For clinicians, the Balance Scale offers a quantitative assessment of movements or functional tasks previously evaluated subjectively. With regular use, clinicians have said that they can more easily describe their patients to colleagues, objectively evaluate the progress or deterioration in their patients, and report those changes in quantitative terms to other team members. Balance Scale scores may also be useful for signalling potential problems or risky behavior when, for example, individuals use insufficient walking aids for their level of ability.

Wider use will provide more evidence on how well it performs in different patient populations. Further tests of clinical usefulness may include the assessment of how well it can aid in prospective clinical decisions such as time of discharge or need for mobility aids. These decisions are based in part on safety considerations, an issue related to balance. The current measurement properties of the Balance Scale are sufficient to encourage use of the instrument in daily practice.

The standards of measurement demonstrated by the Balance Scale also make the instrument appropriate for research. Use of the same measure in clinical practice and research would facilitate an exchange of information and advance our knowledge in the area of balance. Specifically, we need more information on the extent of the problem in various groups and more research on how best to prevent dysfunction, and once impairment of balance occurs, how to retrain it. There are data illustrating the benefits of exercise in the elderly on cardiorespiratory endurance, flexibility, and strength (Hopkins et al 1990; Morey et al 1991). In addition, there is evidence that the elderly can improve with practice, and that overall physical conditioning helps ensure a high level of motor performance in terms of speed and consistency (Light 1990). We need to further examine the amount and type of practise required to prevent dysfunction and improve balance, safe mobility and independent function. This information is important for determining policies on health promotion, allocating services and targeting groups in need of interventions. It is also relevant for falls research to identify individuals at risk of injury, to prevent functional decline, and evaluate interventions that include an exercise component. In addition, falls research requires a measure of balance that can describe the status of the subjects.

At present, it is difficult to compare the findings of studies due to differences in populations, methods and measurements. By definition, a fall is a loss of balance. Thus, a measure of balance must be incorporated into studies researching the risk factors for falls. Balance is, however, not directly related to falls because individuals may compensate for their lack of balance by being more careful, using more aids and restricting their level of activity. This view of falls as an interaction of ability and opportunity to fall is useful for planning interventions.

Use of balance as an outcome measure in intervention studies also offers a more positive approach than assessing the prevention of falls. Interventions can be targeted at increasing ability, offering compensations to improve safety and/or making the environment and level of activity more appropriate to the individual's level of ability. Other variables can be examined as to how they affect either the ability or the opportunity to fall.

An advantage of a performance-based measure is that it allows comparisons across diagnostic categories. This property is particularly useful in the elderly who commonly have multi-system impairments. Additionally, certain diagnoses such as stroke incorporate such a wide range of symptoms and functional deficits that it is impractical to differentiate patients solely on the basis of a diagnosis. Moreover, balance represents a characteristic that can be improved, as demonstrated in stroke patients one year post onset of stroke (Tangeman et al. 1991). In contrast, age or diagnosis cannot be changed.

The Balance Scale appears to have good potential as an outcome measure. As mentioned, a direct comparison with the Tinetti Balance sub-scale showed that the Balance Scale discriminated more efficiently among groups of elderly subjects using different mobility aids. In addition, the only other measure to demonstrate the ability to monitor change in status was the measure of Functional Reach, which is a component item of the Balance Scale.

In conclusion, the Balance Scale has continued to perform well relative to other measures. The results of the present study have contributed to the information on its measurement properties. The final acid test for an instrument is its use as an outcome measure in a clinical trial evaluating the effectiveness of treatments. Such trials are expensive and cannot afford poor choices in measuring instruments. The current knowledge of the properties of the Balance Scale suggests it is ready for this test.

BIBLIOGRAPHY

Ashley MJ, Gryfe CI, Amies A: A longitudinal study of falls in an elderly population. II: Some circumstances of falling. Age Ageing 6: 211-220, 1977

Badke MB, Duncan PW, DiFabio RP: Influence of prior knowledge on automatic and voluntary adjustments in healthy and hemiplegic subjects. Phys Ther 67: 1495-1500, 1987

Badke MB, Duncan PW: Patterns of rapid motor responses during postural adjustments when standing in healthy subjects and hemiplegic patients. Phys Ther 63: 13-20, 1983

Ball K, Sekuler R: Improving visual perception in older observers. J Gerontol 41:176-182, 1986

Basmajian JV, Deluca CJ: Muscles Alive. fifth edition Baltimore, Williams and Wilkens, 1985, 252-264

Belenkii VY, Gurfinkel VS, Paltsev, YI: Element of control of voluntary movements. Biofizika 12: 135-141, 1967

Berg K: Balance and its measure in the elderly: A review. Physiotherapy Canada 41:240-246, 1989

Berg K, Wood-Dauphinee S, Williams JI, Gayton D: Measuring balance in the elderly: preliminary development of an instrument. Physiotherapy Canada 41:304-311, 1989

Berg K, Maki B, Williams JI, Holliday P, Wood-Dauphinee S. A comparison of clinical and laboratory measures of postural balance in an elderly population. Arch Phys Med Rehabil 73: 1073-1080, 1992

Berg K, Wood-Dauphinee S, Williams JI, Maki, B. Measuring balance in the elderly: validation of an instrument. Can. J. Pub. Health 83, supplement 2: S7-11, 1992

Bernstein N: The Coordination and Regulation of Movements. Oxford, Pergamon Press, 1967

Black FO, Wall C III, Rockette HE Jr, Kitch R: Normal subject postural sway during the Romberg test. Am J Otolaryngol 3:309-318, 1982

Black FO, Wall C, Nashner LM: Effects of visual and support surface orientation references upon postural control in vestibular deficient subjects. Acta Otolarynol 95: 199-210, 1983

Bleecker ML, Bolla-Wilson K, Kawas C, Agnew J: Age-specific norms for the Mini-Mental Status Exam. Neurology 38: 1565-1568, 1988

Bobath B: Abnormal Postural Reflex Activity caused by Brain Lesions. London, William Heineman Medical Books Limited, 1965

Bobath B: Adult Hemiplegia:Evaluation and Treatment. second edition, London, William Heineman Medical Books Limited, 1978

Bohannon RW, Larkin PA, Cook AC, Gear J, Singer J: Decrease in timed balance test scores with aging. Phys Ther 64: 1067-1070, 1984

Bohrnstedt GW: Measurement In Handbook of Survey Research. PH Rossi, JD Wright, AB Anderson (eds) New York, Academic Press 1983, 69-121

Bouisset S, Zattara M: Biomechanical study of the programming of anticipatory postural adjustments associated with voluntary movement. J Biomech 20: 735-742, 1987

Bravo G, Potvin L: Estimating the reliability of continuous measures with Cronbach's alpha or the intraclass correlation coefficient: Toward an integration of two traditions. J Clin Epidemiol 44: 381-390, 1991

Briggs RC, Gossman MR, Birch R, Drews JE, Shaddeau SA: Balance performance among noninstitutionalized elderly women. Phys Ther 69: 748-756, 1989

Brocklehurst JC, Robertson D, James-Groom P: Clinical correlates of sway in old age - sensory modalities. Age Ageing 11: 1-10, 1982

Brody EM, Kleban MH, Moss MS, Kleban F: Predictors of falls among institutionalized women with Alzheimer's Disease. J Am Geriatr Soc 32: 877-881, 1984

Brunnstrom S: Movement Therapy in Hemiplegia. New York, Harper and Row, 1970

Campbell AJ, Reinken J, Allan BC, Martinez GS: Falls in old age: a study of frequency and related factors. Age Ageing 10: 264-270, 1981

Campbell AJ, Borrie MJ, Spears GF: Risk factors for falls in a community based prospective study of people 70 years and older. J Gerontol 44: M112-117, 1989

Campbell AJ, Borrie MJ, Spears GF, Jackson SL, Brown JS, Fitzgerald JL: Circumstances and consequences of falls experienced by a community population 70 years and older during a prospective study. Age Ageing 19: 136-141, 1990

Canadian Task Force on the Periodic Health Examination: The Periodic Health

Examination. Can Med Assoc J 121: 1193, 1979.

Carmines EG, Zeller RA: Reliability and Validity Assessment. Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-017. Beverly Hills, Sage Publications, 1979.

Carrol D: Disability in hemiplegia caused by cerebrovascular disease: serial study of 98 cases. J Chron Dis 15:179-185, 1962.

Chandler JM, Duncan PW, Studenski SA: Balance performance on the postural stress test: comparison of young adults, healthy elderly and fallers. Phys Ther 70: 410-415, 1990

Chino N: Efficacy of Barthel Index in evaluating activites of daily living in Japan, the United States and United Kingdom. Stroke 21 (Suppl II): II 64-65, 1990

Clarke B, Gowland C, Brandstater M, deBruin H: A re-evaluation of the Brunnstrom assessment of motor recovery of the lower limb. Physiother Can 35: 207-211, 1983.

Cohen J: Statistical Power Analysis for the Behavioral Sciences. New York, Academic Press 1977, 8-17, 274-288

Cohn TE, Lasley DJ: Visual depth illusion and falls in the elderly. Clin Geriatr Med 1: 601-620, 1985

Colton T: Statistics in Medicine. Boston, Little, Brown and Company, 1974.

Cordo PJ, Nashner LM: Properties of postural adjustments associated with rapid arm movements. J Neurophysiol 47: 287-302, 1982

Crilly RG, Willems DA, Trenholme KJ, Hayes KC, Delaquerriere- Richardson LFO: Effect of exercise on postural sway in the elderly. Gerontol 35: 137-143, 1989

Cronbach LJ: Coefficient alpha and the internal structure of test. Psychometrika 16:297-334, 1951

Cummings SR, Nevitt MC, Kidd S: Forgetting falls: The limited accuracy of recall of falls in the elderly. J Amer Geriatr Soc 36: 615-616, 1988

Dettmann MA, Linder MT, Sepic SB: Relationships among walking performance, postural stability, assessments of the hemiplegic patient. Amer J Phys Med 66: 77-90, 1987

Deyo RA, Inui TS: Toward clinical applications of health status measures: sensitivity of scales to clinically important changes. Health Services Research 19: 275-289, 1984

Deyo R: Measuring functional outcomes in therapeutic trials for chronic disease. Controlled Clinical Trials 5:223-240, 1984

Deyo RA, Centor RM: Assessing the responsiveness of functional scales to clinical change: an analogy to diagnostic test performance. J Chron Dis 39: 897-906, 1986

Deyo RA: Measuring the functional status of patients with low back pain. Arch Phys Med Rehabil 69: 1044-1052, 1988

Deyo RA, Diehr P, Patrick DL: Reproducibility and responsiveness of health status measures: statistics and strategies for evaluation. Controlled Clinical Trials 12: 142S-158S, 1991

Donaldson SW, Wagner CC, Gresham GE: Unified ADL evaluation form. Arch Phys Med Rehabil 54: 175-179, 1973.

Doman J, Fernie GR, Holliday PJ: Visual input: Its importance in the control of postural sway. Arch Phys Med Rehabil 59: 586-591, 1978

Drowatzky JN, Zuccato FC: Interrelationships between selected measures of static and dynamic balance. Research Q 38:509-510, 1967

Duncan PW, Propst M, Nelson SG: Reliability of the Fugl-Meyer assessment of sensorimotor recovery following cerebrovascular accident. Phys Ther 63: 1606-1610, 1983

Duncan PW, Studenski S, Chandler J, Bloomfeld R, Lapointe LK: Electromyographic analysis of postural adjustments in two methods of balance testing. Phys Ther 70: 88-96, 1990

Duncan PW, Weiner DK, Chandler J, Studenski S: Functional reach: A new clinical measure of balance. J Gerontol 45: M192-197, 1990

Duncan PW, Studenski S, Chandler J, Prescott B: Functional Reach: Predictive validity in a sample of elderly male veterans. J Gerontol 47: M93-98, 1992

Ebel RL: Estimation of the reliability of ratings. Psychometrika 16:407-424, 1951

Eggert GM, Granger CV, Morris R et al: Caring for patients with long term disability. Geriatrics 32: 102-114, 1977

Era P, Heikkinen E: Postural sway during standing and unexpected disturbance of balance in random samples of men of different ages. J Gerontol 40: 287-295, 1985

Feinstein AR, Josephy BR, Wells CK: Scientific and clinical problems in indexes of functional disability. Ann Intern Med 105: 413-420, 1986

Feinstein AR: Clinimetrics, New Haven, Yale University Press, 1987, 6-21

Felson DT, Anderson JJ, Hannan MT, Milton RC, Wilson PWF, Kiel DP: Impaired vision and hip fracture: The Framingham study. J Amer Geriatr Soc 37: 495-500, 1989

Fernie GR, Holliday PJ: Postural sway in amputees and normal subjects. J Bone and Joint Surg 60A: 895-898, 1978

Fernie GR, Gryfe CI, Holliday PJ, Llewellyn A: The relationship of postural sway in standing to the incidence of falls in geriatric subjects. Age and Ageing 11: 11-16, 1982

Fiatorone MA, Evans WJ: Exercise in the oldest old. Top Geriatr Rehabil 5:63-77, 1990

Fleiss JL: The Design and Analysis of Clinical Experiments. Toronto, John Wiley and Sons, 1986, 1-32

Folstein MF, Folstein SE, McHugh PR: Mini Mental State a practical method for grading the cognitive state of patients for the clinician. J Psychiat Res 12:189-198, 1975

Frank JS, Patla AE, Brown J: Characteristics of postural control accompanying voluntary arm movement in the elderly. Soc Neurosci Abstracts 13: 335, 1987

Fugl-Meyer AR, Jaaski L, Legman I, Olsson S, Steglind S: The poststroke hemiplegic patient. 1. a method for evaluation of physical performance. Scand J Rehab Med 7: 13-31, 1975

Fugl-Meyer AR: The effect of rehabilitation in hemiplegia as reflected in the relation between motor recovery and ADL function. Proceedings IRMA II, Mexico City, 1976, Abstract p683

Gabell A, Simons MB: Balance coding. Physiotherapy 68: 286-288, 1982

Gabell A, Simons MA, Nayak USL: Falls in the healthy elderly: predisposing causes. Ergonomics 28: 965-975, 1985

Gabell A: Falls in the elderly: Will dance reduce their incidence? J Hum Movement studies 12: 119- 132, 1986

Gahery Y: Associated movements, postural adjustments and synergie: Some comments about the history and significance of three motor concepts. Arch Ital Biol 125: 345-360, 1987

Galley PM, Forster AL: Human Movement: An Introductory Text for Physiotherapy Students. second edition Melbourne, Churchill Livingstone, 1987, 86-97

Gauthier-Gagnon C, St. Pierre D, Drouin G, Riley E: Augmented sensory feedback in the early training of standing balance of below-knee amputees. Physiother Can 38: 137-142, 1986

Gehlsen GM, Whaley MH: Falls in the elderly: Part I Gait. Arch Phys Med Rehabil 71:735-738, 1990.

Gehlsen GM, Whaley MH: Falls in the elderly: Part II, Balance, strength and flexibility. Arch Phys Med Rehabil 71:739-741, 1990.

Goldie PA, Bach TM, Evans OM: Force platform measures for evaluating postural control: Reliability and validity. Arch Phys Med Rehabil 70: 510-517, 1989

Goldie PA, Matyas TA, Spencer KI, McGinley RB: Postural control in standing following stroke: Test-retest reliability of some quanitative clinical tests. Phys Ther 70: 234-243, 1990

Granger CV, Greer DS, Liset E et al: Measurement of outcomes of care for stroke patients. Stroke 6: 34-41, 1975

Granger CV, Greer DS: Functional status measurement and medical rehabiliation outcomes. Arch Phys Med Rehabil 57: 103-109, 1976

Granger CV, Sherwood CC, Greer DS: Functional status measures in comprehensive stroke care program. Arch Phys Med Rehabil 58: 555-561, 1977

Granger CV, Albrecht GL, Hamilton BB: Outcome of comprehensive medical rehabilitation: measurement by Pulse Profile and the Barthel Index. Arch Phys Med Rehabil 60: 145-154, 1979

Granger CV, Deurs LS, Peters NC et al: Stroke rehabiliation: analysis of repeated Barthel Index measures. Arch Phys Med Rehabil 60: 14-17, 1979

Granger CV, Hamilton BB, Gresham GE: The stroke rehabilitation outcome study: Part I. General Description. Arch Phys Med Rehabil 69: 506-509, 1988

Granger CV, Hamilton BB, Gresham GE, Kramer AA: The stroke rehabilitation outcome study: Part II. Relative merits of the total Barthel Index score and a four-item subscore in predicting patient outcomes. Arch Phys Med Rehabil 70:100-103, 1989

Gresham GE, Fitzpatrick TE, Wolf PA et al: Residual disability in survivors of stroke - the Framingham study. N Engl J Med 293: 954-956, 1975

Gresham GE, Phillips TF, Labi MLC: ADL status in stroke: relative merits of three standard indexes. Arch Phys Med Rehabil 61: 355-358, 1980

Grimby G: Muscle changes and trainability in the elderly. Top Geriatr Rehabil 5: 54-62, 1990

Grisso JA, Kelsey JL, Strom BL, Chiu GY, Maislin G, O'Brien LA, Hoffman S, Kaplan F, and the Northeast Fracture Group: Risk factors for falls as a cause of hip fracture in women. 324: 1326-1331, 1991

Gryfe CI, Amies A, Ashley MJ: A longitudinal study of falls in an elderly population l Incidence and morbidity. Age Ageing 6: 201-210, 1977

Guiliani CA: Theories of motor control: New concepts for physical therapy. In Contemporary Management of Motor Control Problems. Proceedings of the II STEP Conference. Foundation for Physical Therapy, Alexandria, Virginia, 1991, 29-35

Guyatt G, Walter S, Norman G: Measuring change over time: assessing the usefulness of evaluative instruments J Chron Dis 40: 171-178, 1987

Guyatt G, Deyo RA, Charlson M, Levine MN, Mitchell A: Responsiveness and validity in health status measurement: A clarification. J Clin Epidemiol 42: 403-408, 1989

Hasselkus BR, Shambes GM: Aging and postural sway in women. J Gerontol 30: 661-667, 1975

Hayes KC, Spencer CL, Riach CL, Lucy SD, Kirshen AJ: Age related changes in postural sway. In Biomechanics IX-A. DA Winter, RW Norman, RP Wells, KC Hayes, AE Patla (eds) Champagne, Illnois, Human Kinetics Publishers Inc, 1985, 383-387

Hayes KC, Riach CL: Preparatory postural adjustments and postural sway in young children. In Development of Posture and Gait Across the Life Span. MH Woollacott, A Shumway-Cook (eds) Columbia, South Carolina, University of South Carolina Press, 1989, 97-127

Hays RD, Hadorn D: Responsiveness to change: An aspect of validity, not a separate dimension. Quality of Life Research 1: 73-75, 1992

Helmstadter GC: Principles of Psychological Measurement. New York, Appleton-Century-Crofts, 1964

Hertanu JS, Demopoulos JT, Yang WC, Calhoun WF, Fenigstein HA: Stroke rehabilitation: correlation and prognostic value of computerized tomography and sequential functional assessments. Arch Phys Med and Rehabil 65: 505-508, 1984

Hewer RL: Outcome measures in stroke: A British view. Stroke 21 (Suppl II):II 52-55, 1990

Higgins S: Motor skill acquisition. Movement Science Series, Part 3, Phys Ther 71: 123-139, 1991

Hill K, Vandervoort A, Kramer J: Reproducibility of performances on a test of postural responses in healthy elderly females. Physiother Can 42: 61-67, 1990

Hocherman S, Dickstein R, Pillar T: Platform training and postural stability in hemiplegia. Arch Phys Med Rehabil 65: 588-592,1984

Holliday PJ, Fernie GR: Changes in the measurement of postural sway resulting from repeated testing. Aggressologie 20: 225-228, 1979

Hopkins DR, Murrah B, Hoeger WWK, Rhodes RC: Effect of low impact aerobic dance on the functional fitness of elderly women. The Gerontologist 30: 189-192, 1990

Horak FB, Esselman P, Anderson ME, Lynch MK: The effects of movement velocity, mass displaced, and task certainty on associated postural adjustments made by normal and hemiplegic individuals. J Neurol Neurosurg Psychiatry 47: 1020- 1028, 1984

Horak FB, Nashner LM: Central programming of postural control: adaptation to altered support surface configurations. J Neurophysiol 55: 1369-1381, 1986

Horak FB: Clinical measurement of postural control in adults. Phys Ther 67: 1881-1885, 1987

Horak FB, Mirka A, Schupert CL: The role of peripheral vestibular disorders in postural dyscontrol in the elderly. In Development of Posture and Gait Across the Life Span. MH Woollacott, A Shumway-Cook (eds) Columbia, South Carolina, University of South Carolina Press, 1989, 253-279

Horak FB, Shupert CL, Mirka A: Components of postural dyscontrol in the elderly: A review. Neurobiol Aging 10: 727-738, 1989

Horak FB: Assumptions underlying motor control for neurologic rehabilitation. In Contemporary Management of Motor Control Problems. Proceedings of the II STEP Conference. Foundation for Physical Therapy, Alexandria, Virginia, 1991, 11-27

Inglin B, Woollacott MH: Age-related changes in anticipatory postural adjustments associated with arm movements. J Gerontol 43: M105-113, 1988

Isaacs B: Clinical and laboratory studies of falls in old people. Clin Geriatr Med 1: 513-524, 1985

Jaeschke R, Singer J, Guyatt GH: Measurement of health status: Ascertaining the minimal clinically important difference. Controlled Clinical Trials 10: 407-415, 1989

Jebsen RH, Trieschmann RB, Mikulic MA, Hartley RB, McMillan J, Snook ME: Measurement of time in a standardized test of patient mobility Arch Phys Med Rehabil 53: 170-175, 1970

Jette AM: Concepts of health and methodological issues in functional assessment. In: Functional Assessment in Rehabilitation Medicine. CV Granger and GE Gresham (eds) Baltimore, Williams and Wilkins, 1984, 46-64

Jette AM, Branch LG, Berlin J: Musculoskeletal impairments and physical disablement among the aged. J Gerontol 45: M203-208, 1990

Kane RA, Kane RL: Assessing the Elderly: A practical guide to measurement. Toronto, Lexington Books, 1981, 1-23,247-271

Kauffman T: Impact of aging-related musculoskeletal and postural changes on falls. Top Geriatr Rehabil 5: 34-43, 1990

Kazis LE, Anderson JJ, Meenan RF: Effect sizes for interpreting changes in health status. Medical Care 27 (Suppl): S178-189, 1989

Kellogg International Work Group: The prevention of falls in later life. Dan Med Bull 34 (Suppl 4):1-24, 1987

Kelsey JL, Thompson WD, Evans AS: Methods in Observational Epidemiology. New York, Oxford University Press 1986

Kerlinger FN: Foundations of Behavioural Research. Third edition Toronto, Holt, Rinehart and Winston Inc, 1986

Keshner EA: How theoretical framework biases evaluation and treatment. In Contemporary Management of Motor Control Problems. Proceedings of the II STEP

Conference. Foundation for Physical Therapy, Alexandria, Virginia, 1991, 37-47

Kirshner B, Guyatt G: A methodological framework for assessing health indices. J Chron Dis 38: 27-36, 1985

Kleinbaum D, Kupper L,Muller KE: Applied Regression Analysis and other Multivariable Methods. Boston, PWS-Kent Publishing Company, 1988

Kusoffsky A, Wadell I, Nilsson BY: The relationship between sensory impairment and motor recovery in patients with hemiplegia. Scand J Rehab Med 14:27-32, 1982

Last JM. editor. A Dictionary of Epidemiology second edition New York, Oxford University Press, 1988

Lee WA, Deming L, Sahgal V: Quantitative and clinical measures of static standing balance in hemiparetic and normal subjects. Phys Ther 68: 970-976, 1988

Leibowitz HW, Shupert CL: Spatial orientation mechanisms and their implications for falls. Clin Geriatr Med 1: 571-580, 1985

Liang MH, Larson MG, Cullen KE, Schwartz JA: Comparative measurement efficiency and sensitivity of five health status instruments for arthritis research. Arthritis Rheum 28: 542-547, 1985

Liang MH, Fossel AH, Larson MG: Comparisons of five health status instruments for orthopaedic evaluation. Med Care 28: 632-642, 1990

Lichenstein MJ, Shields SL, Shiavi RG, Burger C: Clinical determinants of biomechanics platform measures of balance in aged women. J Am Geriatr Soc 36: 996-1002, 1988

Lichenstein MJ, Shields SL, Shiavi RG, Burger C: Exercise and balance in aged women: a pilot controlled clinical trial. Arch Phys Med Rehabil 70: 138-143, 1989

Lichenstein MJ, Burger C, Shields SL, Shiavi RG: Comparison of biomechanics platform measures of balance and videotaped measures of gait with a clinical mobility scale in elderly women. J Gerontol 45: M49-54, 1990

Light K: Information processing for motor performance in aging adults. Phys Ther 70: 820-826, 1990

Lipshitz LA: Abnormalities in blood pressure homeostasis that contribute to falls in the elderly. Clin Geriatr Med 1: 637-645, 1985

Lipshitz LA, Jonsson PV, Kelley MM, Koestner JS: Causes and correlates of recurrent falls in ambulatory frail elderly: 46: M114-122, 1991

Macdonald JB: The role of drugs in falls in the elderly. Clin Geriatr Med 1:621-632, 1985

MacKenzie CR, Charlson ME, DiGiola D, Kelley K: Can the sickness impact profile measure change? An example of scale assessment. J Chron Dis 39: 429-438, 1986

Magnus R: Some results of studies in the physiology of posture. Lancet 211: 531-536 and 585-588, 1926

Mahoney FI, Barthel DW: Functional evaluation: The Barthel Index Md State Med J 14:61-65, 1965

Maki BE, Holliday PJ, Fernie GR: A posture control model and balance test for the prediction of relative postural stability. IEEE Trans Biomed Eng BME 34:797-810, 1987

Maki BE, Fernie GR: A system identification approach to balance testing. Prog Brain Res 76: 297-306, 1988

Maki BE, Holliday PJ, Fernie GR: Aging and postural control: A comparison of spontaneous and induced-sway balance tests. J Amer Geriatr Soc 38: 1-9, 1990

Maki BE, Holliday PJ, Topper AK: The fear of falling and postural performance in the elderly. J Gerontol 46:M123-131, 1991

Manchester D, Woollacott M, Zederbauer-Hylton N, Marin O: Visual, vestibular and somatosensory contributions to balance control in the older adult. J Gerontol 44: M118-127, 1989

Mathias S, Nayak USL, Isaacs B: Balance in elderly patients: The "Get-up and Go" test. Arch Phys Med Rehabil 67: 387-389, 1986

Mayo NE, Korner-Bitensky N, Becker R, Georges P: Predicting falls among patients in a rehabilitation hospital. Amer J Phys Med Rehab 68: 139-146, 1989

Mayo NE, Korner-Bitensky N, Becker R, Georges P: Preventing falls among patients in a rehabilitation hospital. Can J Rehab 2:235-240, 1989

Mayo NE, Korner-Bitensky N, Kaizer F: Relationship between response time and falls among stroke patients undergoing physical rehabilitation. Int J of Rehab Res 13: 47-55, 1990

McDowell I, Newell C: Measuring Health: A Guide to Rating Scales and Questionnaires. New York, Oxford University Press, 1987, 3-54

Meenan RF, Anderson JJ, Kazis LE, Egger MJ, Altz-Smith M, Samuelson CO Jr, Wilkens RF, Solsky MA, Hayes SP, Blocka KI, Weinstein A, Guttadauria M, Kaplan SB, Klippel J: Outcome assessment in clinical trials : Evidence for the sensitivity of a health status measure. Arthritis Rheum 27:1344-1352, 1984

Mikulic MA, Griffith ER, Jebsen RH: Clinical applications of a standardized mobility test. Arch Phys Med Rehabil 57: 143-146, 1976

Morey MC, Cowper PA, Feussner JR, Dipasquale RC, Crowley GM, Sullivan RJ: Two-year trends in physical performance following supervised exercise among community dwelling older veterans. J Am Geriatr Soc 39: 549-554, 1991

Morse JM: Computerized evaluation of a scale to identify the fall-prone patient. Can J Public Health 77: 21--25, 1986

Morse JM, Tylko SJ, Dixon HA: Characteristics of the fall prone patient. The Gerontologist 27: 516-522, 1987

Morse JM, Morse RM, Tylko SJ: Development of a scale to identify the fall-prone patient. Can J Aging 8: 366-377, 1989

Mossey JM: Social and psychological factors related to falls among the elderly. Clin Geriatr Med 1: 541-553, 1985

Murray MP, Seireg AA, Sepic SB: Normal postural stability and steadiness: Quantitative assessment. J Bone Joint Surg 57A: 510-515, 1975

Nashner LM: A model describing vestibular detction of body sway motion. Acta Otolaryng 72: 429-436, 1971

Nashner LM: Adapting reflexes controlling the human posture. Exp Brain Res 26: 59-72, 1976

Nashner LM: Fixed patterns of rapid postural responses among leg muscles during stance. Exp Brain Res 30: 13-24, 1977

Nashner LM, Woollacott M, Tuma G: Organization of rapid responses to postural and locomotor-like perturbations of standing man. Exp Brain Res 36, 463-476, 1979

Nashner LM: Balance adjustments of humans perturbed while walking. J Neurophysiol 44: 650-664, 1980

Nashner LM, Cordo PJ: Relation of automatic postural responses and reaction-time voluntary movements of human leg muscles. Exp Brain Res 43: 395-405, 1981

Nashner LM, Black FO, Wall C: Adaptation to altered support surfaces and visual conditions during stance: Patients with vestibular deficits. J Neuroscience 2: 536-544, 1982

Nashner LM, Diener HC, Horak FB: Selection of human postural synergies differs with peripheral somatosensory versus vestibular loss. In: The proceedings of the Society for Neuroscience fifteenth Annual Meeting 1985

Nashner LM, McCollum G: The organization of human postural movements: A formal basis and experimental synthesis. The Behavioral and Brain Sciences 8: 135-172, 1985

Nashner LM: A systems approach to understanding and assessing orientation and balance disorders. Clackmas, OR NeuroCom International, Inc 1987

Nayak USL, Gabell A, Simons MA, Isaacs B: Measurement of gait and balance in the elderly. J Am Geriatr Soc 30: 516-520, 1982

Nelson AJ: Functional ambulation profile. Phys Ther 54: 1059-1065, 1974

Nickens H: Intrinsic factors in falling among the elderly. Arch Intern Med 145:1089-1093, 1985

Norman GR: Issues in the use of change scores in randomized trials. J Clin Epidemiol 42: 1097-1105, 1989

Norusis MJ/ SPSS Inc. SPSS/PC+ Advanced Statistics V2.0 Chicago, SPSS Inc. 1988

Norusis MJ/ SPSS Inc. SPSS/PC+ Advanced Statistics 4.0 Chicago, SPSS Inc. 1990

Nunnally JC: Psychometric Theory. second edition, New York, McGraw-Hill, 1978

Overstall PW, Exton-Smith AN, Imms FJ, Johnson AL: Falls in the elderly related to postural imbalance. Brit Med J 1: 261-264, 1977

Overstall PW: Prevention of falls in the elderly. J Am Geriatr Soc 26: 481-484, 1980

Overstall PW, Hazell JWP, Johnson AL: Vertigo in the elderly. Age and Ageing 10: 105-109, 1982

Patla A, Frank J, Winter D: Assessment of balance control in the elderly: major issues. Physiother Can 42: 89-97, 1990

Peck WA: Falls and hip fracture in the elderly. Hospital Practice 72: A-L, 1986

Podsiadlo D, Richardson S: The Timed "Up and Go" a test of basic mobility for frail elderly persons. J Am Geriatr Soc 39: 142-148, 1991

Polit A, Bizzi E: Characteristics of motor programs underlying arm movements in monkeys. J Neurophysiol 42: 183-194, 1979

Potvin AR, Tourtellotte WW: The neurological examination: advancements in its quantification. Arch Phys Med Rehabil 56: 425-437, 1975

Potvin AR, Syndulko K, Tourtellotte WW, Lemmon JA, Potvin JH: Human neurologic function and the aging process. J Am Geriatr Soc 28: 1-9, 1980

Prudham D, Grimley Evans J: Factors associated with fals in the elderly: a community study. Age Ageing 10: 141-146, 1981

Raab DM, Agre JC, McAdam M, Smith EL: Light resistence and stretching exercise in elderly woemn: Effect upon flexibility. Arch Phys Med Rehabil 69: 268-272, 1988

Ray WA, Griffin MR: Prescribed medications and the risk of falling. Top Geriatr Rehabil 5: 12-20, 1990

Reding MJ: A model stroke classification scheme and its use in outcome research. Stroke 21 (Suppl II): II 35-37, 1990

Reed ES: An outline of a theory of action systems. J Motor Behav 14: 98-104, 1982

Reed ES: Changing theories of postural development. In Development of Posture and Gait Across the Life Span. MH Woollacott, A Shumway-Cook (eds) Columbia, South Carolina, University of South Carolina Press, 1989, 3-24

Ring C, Matthews R, Nayak L, Isaacs B: Visual Push: A sensitive measure of dynamic balance in man. Arch Phys Med Rehabil 69: 256-260, 1988

Ring C, Nayak L, Isaacs B: Balance function in elderly people who have and have not fallen. Arch Phys Med Rehabil 69: 261-264, 1988

Rosenhall U: Degenerative patterns in the aging human vestibular neuro-epithelia. Acta Otolaryng 76: 208-220, 1973

Rosenhall U, Rubin W: Degenerative changes in the human vestibular sensory epithelia. Acta Otolaryng 79: 67-80, 1975

Rusin MJ: Stroke rehabilitation: a geropsychological perspective. Arch Phys Med Rehabil 71:914-922, 1990

Schulmann DL, Godfrey B, Fisher AG: Effect of eye movements on dynamic equilibrium. Phys Ther 67:1054-1057, 1987

Seliktar R, susak Z, Najenson T, Solzi P: Dynamic features of standing their correlation with neurological disorders. Scand J Rehabil Med 10: 59-64, 1978

Sheldon JH: Effect of age on control of sway. Geront Clin 5: 129-138, 1963

Sherrington CS: The Integrative Action of the Central Nervous System 1947 edition New York, Cambridge University Press, 1947

Shumway-Cook A, Woollacott MH: The growth of stability: postural control from a developmental perspective. J Motor Behav 17: 131-147, 1985

Shumway-Cook A, Horak FB: Assessing the influence of sensory interaction on balance. Suggestion from the field. Phys Ther 66: 1548-1550, 1986

Shumway-Cook A, McCollum G: Assessment and treatment of balance deficits In Motor Control and Physical Therapy: Theoretical Framework and Practical Applications PC Montgomery, BH Connolly (eds) Hixson TN Chattanooga Group, Inc 1991, 123-138

Spector PE: Summated rating scale construction : An introduction. Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-082. Newbury Park, Sage Publications 1992

Spitzer WO, Dobson AJ, Hall J, Chesterman E, Levi J, Battista RN, Catchlove BR: Measuring the quality of life of cancer patients. J Chron Dis 34: 585-597, 1981

Stevens S: Mathematics, measurement and psychophysics. In Handbook of Experimental Psychology. New York, John Wiley and Sons Inc, 1951

Stelmach GE, Worrington CJ: Sensorimotor deficits related to postural instability: Implications for falling in the elderly. Clin Geriatr Med 1: 679-694, 1985

Stelmach GE, Zelaznik MC, Lowe D: The influence of aging and attentional demands on recovery from postural instability. Aging 2: 155-161, 1990

Streiner DL, Norman GR: Health Measurement Scales: A Practical Guide to their Development and Use. New York, Oxford University Press, 1989

Studenski S, Duncan PW, Chandler J: Postural responses and effector factors in persons with unexplained falls: Results and methodologic issues. J Am Geriatr Soc 39: 229-234, 1991

Swift CG: Postural instability as a measure of sedative drug response. Br J Clin Pharmac 18: 87S-90S, 1984

Tangeman PT, Banaitis DA, Williams AK: Rehabilitation of chronic stroke patients: changes in functional performance. Arch Phys Med Rehabil 71:876-880, 1990

Thelen E, Kelso JAS, Fogel A: Self-organizing systems and infant motor development. Developmental Review 7: 39-65, 1987

Thyssen HH, Brynskov J, Jansen EC, Munster-Swendsen J: Normal ranges and reproducibility for the quantitative Romberg's test. Acta Neurol. Scandinav. 66: 100-104, 1982

Tinetti ME: Performance-oriented assessment of mobility problems in elderly patients. J Am Geriatr Soc 34:119-126, 1986

Tinetti ME, Williams TF, Mayewski R: A fall risk index for elderly patients based on number of chronic disabilities. Am J Med 80: 429-434, 1986

Tinetti ME: Factors associated with serious injury during falls by ambulatory nursing home residents. J Am Geriatr Soc 35: 644-648, 1987

Tinetti ME, Speechley M, Ginter SF: Risk factors for falls among elderly persons living in the community. N Eng J Med 319: 1701-1707, 1988

Tinetti ME, Ginter SF: Identifying mobility dysfunctions in elderly patients: standard neurological examination or direct assessment. JAMA 259: 1190-1193, 1988

Tinetti ME, Richman D, Powell L: Falls efficacy as a measure of fear of falling. J Gerontol 45:P239-243, 1990

Tobis JS, Reinsch S, Swanson JM, Byrd M, Scharf T: Visual perception dominance of fallers among community-dwelling older adults. J Am Geriatr Soc 33: 330-333, 1985

Tombaugh TN, McIntyre NJ: The Mini-Mental State Examination: A comprehensive review. J Am Geriatr Soc 40: 922-935, 1992

Traub M, Rothwell JC, Marsden CD: Anticipatory postural reflexes in Parkinson's Disease and other akinetic-rigid syndromes and in cerebellar ataxia. Brain 103: 393-412, 1980

Tuley MR, Mulrow CD, McMahan CA: Estimating and testing an index of responsiveness and the relationship of the index to power. J Clin Epidemiol 44: 417-421, 1991

Uden G: Inpatient accidents in hospitals. J Am Geriatr Soc 33: 833-841, 1985

Vlahov D, Myers A, Al-Ibrahim MS: Epidemiology of falls among patients in a rehabilitation hospital. Arch Phys Med Rehabil 71: 8-12, 1990

Wade DT: Measurement in Neurological Rehabilitation Oxford, The Oxford University Press, 1992, 70-76

Wade DT, Collin C: The Barthel ADL Index: A standard measure of physical disability? Int Disab Stud 10: 64-67, 1988

Warren WH, Blackwell AW, Morris MW: Age differences in perceiving the direction of self-motion from optical flow. J Gerontol 44: P147-153, 1989

Weiner DK, Bongoneri D, Kochersberger G, Duncan PW, Studenski S: Changes in functional reach during rehabilitation. (Abstract) J Am Geriatr Soc 39:A53, 1991

Weiner WJ, Nora LM, Glantz RH: Elderly inpatients: postural reflex impairment. Neurology 34: 945-947, 1984

Whanger AD, Wang HS: Clinical correlates in elderly psychiatric patients. J Gerontol 29: 39-45, 1974

Whipple RH, Wolfson LI, Amerman PM: The relationship of knee and ankle weakness to falls in nursing home residents: an isokinetic study. J Am Geriatr Soc 35: 13-20, 1987

Wild D, Nayak USL, Isaacs B: Prognosis of falls in old people at home. J Epidemiol Community Health 35: 200-204, 1981

Wild D, Nayak USL, Isaacs B: How dangerous are falls in old people at home? Brit Med J 282: 266-268, 1981

Wild D, Nayak USL, Isaacs B: Description, classification and prevention of falls in old people at home. Rheumatol Rehabil 20: 153-159, 1981

Winstein CJ, Gardner ER, McNeal DR, Barto PS, Nicholson DE: Standing balance training: Effect on balance and motion in hemiparetic adults. Arch Phys Med Rehabil 70: 755-761,1989

Wolfson LI, Whipple RH, Amerman P, Kaplan J, Kleinberg A: Gait and balance in the elderly. Two functional capacities that link sensory and motor ability to falls. Clin Geriatr Med 1: 649-659, 1985

Wolfson LI, Whipple RH, Amerman P, Kleinberg A: Stressing the postural response: A quantitative method for testing balance J Am Geriatr Soc 34: 845-850, 1986

Wood-Dauphinee S: A trial of team care in the treatment of acute stroke. Stroke 15: 864-872, 1984

Wood-Dauphinee S, Williams JI, Opzoomer A: Reintegration to normal living as a proxy to quality of life. J Chron Dis 40:491-499, 1987

Wood-Dauphinee SL, Williams JI, Shapiro SH: Examining outcome measures in a clinical study of stroke. Stroke 21:731-739, 1990

Woollacott MH, Shumway-Cook A, Nashner L: Postural reflexes and aging. In: The Aging Motor System. ed FJ Pirozzolo and GJ Maletta, Praeger, New York, 1982, 98-119

Woollacott MH: Gait and postural control in the aging adult. In: Disorders of Posture and Gait. W. Bles and T. Brandt (eds) Elsevier Science Publishers, The Netherlands, 1986, 325-336

Woollacott MH, Shumway-Cook A, NAshner LM: Aging and posture control: Changes in sensory organization and muscular coordination. Int J Aging Hum Develop 23: 97-114, 1986

Woollacott MH: Changes in posture and voluntary control in the elderly: Research findings and rehabilitation. Top Geriatr Rehabil 5: 1-11, 1990

Wylie CM: The value of early rehabilitation in stroke. Geriatrics 25:107-113, 1970

APPENDIX 1

BALANCE VALIDATION STUDY
BALANCE SCALE WORK SHEET

Name of subject:

Study number:

Location:

Room number:

Date:

Name of Rater:

Evaluation: initial__ 2 __ 3 __ 4 __

Reliability : inter __ intra __

ITEM	DESCRIPTION	SCORE (0-4)
1	Sitting to standing	__
2	Standing unsupported	__
3	Sitting unsupported	__
4	Standing to sitting	__
5	Transfers	__
6	Standing with eyes closed	__
7	Standing with feet together	__
8	Reaching forward with outstretched arm	__
9	Retrieving object from floor	__
10	Turning to look behind	__
11	Turning 360 degrees	__
12	Placing alternate foot on stool	__
13	Standing with one foot in front	__
14	Standing on one foot	__

BALANCE SCALE developed in partial fulfilment of Master of Science degree McGill University: K Berg 1988

1 SITTING TO STANDING

INSTRUCTION: Please stand up. Try not to use your hands for support.

GRADING: Please mark the lowest category which applies.

( )	( )	( )	( )	( )	( )
4	3	2	1	0	
able to stand no hands and stabilize indep	able to stand indep using hands	able to stand using hands after several tries	needs minimal aid to stand or to stabilize	needs moderate or maximal assist to stand	

2 STANDING UNSUPPORTED

INSTRUCTION: Stand for two minutes without holding.

GRADING: Please mark the lowest category which applies.

( )	( )	( )	( )	( )	
4	3	2	1	0	( )
able to stand safely 2 min	able to stand 2 min with supervision	able to stand 30 seconds unsupported	needs several tries to stand 30 sec unsupported	unable to stand 30 sec unassisted	

IF SUBJECT ABLE TO STAND 2 MIN SAFELY, SCORE FULL MARKS FOR SITTING UNSUPPORTED. PROCEED TO POSITION CHANGE STANDING TO SITTING.

3 Sitting unsupported feet on floor

INSTRUCTION: Sit with arms folded for two minutes.

GRADING: Please mark the lowest category which applies.

( )	( )	( )	( )	( )	
4	3	2	1	0	( )
able to sit safely and securely 2 minutes	able to sit 2 minutes under supervision	able to sit 30 seconds	able to sit 10 seconds	unable to sit without support 10 sec	

4 STANDING TO SITTING

INSTRUCTION: Please sit down.

GRADING: Please mark the lowest category which applies.

( )	( )	( )	( )	( )	
4	3	2	1	0	( )
sits safely with minimal use of hands	controls descent by using hands	uses back of legs against chair to control descent	sits indep but has uncontrolled descent	needs assistance to sit	

5 TRANSFERS

INSTRUCTION: Please move from chair to bed and back again. One way toward a seat with armrests and one way toward a seat without armrests.

GRADING: Please mark the lowest category which applies.

( )	( )	( )	( )	( )	
4	3	2	1	0	( )
able to transfer safely with minor use of hands	able to transfer safely definite need of hands	able to transfer with verbal cuing and/or supervision	needs one person to assist	needs two people to assist or supervise to be safe	

6 STANDING UNSUPPORTED WITH EYES CLOSED

INSTRUCTION: Close your eyes and stand still for 10 seconds.

GRADING: Please mark the lowest category which applies.

( )	( )	( )	( )	( )	
4	3	2	1	0	( )
able to stand 10 sec safely	able to stand 10 sec with supervision	able to stand 3 sec	unable to keep eyes closed 3 sec but stays steady	needs help to keep from falling	

7 STANDING UNSUPPORTED WITH FEET TOGETHER

INSTRUCTION: Place your feet together and stand without holding.

GRADING: Please mark the lowest category which applies.

( )	( )	( )	( )	( )	
4	3	2	1	0	( )
able to place feet together indep and stand 1 min safely	able to place feet together indep and stand for 1 min with supervision	able to place feet together indep but unable to hold for 30 sec	needs help to attain position but able to stand 15 sec feet together	needs help to attain position and unable to hold for 15 sec	

8 REACHING FORWARD WITH OUTSTRETCHED ARM WHILE STANDING

INSTRUCTIONS: Lift arm to 90 degrees. Stretch out your fingers and reach forward as far as you can. (Examiner places a ruler at end of fingertips when arm is at 90 degrees. Fingers should not touch the ruler while reaching forward. The recorded measure is the distance forward that the fingers reach while the subject is in the most forward lean position.)

GRADING: Please mark the lowest category which applies.

( )	( )	( )	( )	( )	
4	3	2	1	0	( )
can reach forward confidently >10 inches	can reach forward >5 inches safely	can reach forward >2 inches safely	reaches forward but needs supervision	needs help to keep from falling	

9 PICK UP OBJECT FROM THE FLOOR FROM A STANDING POSITION

INSTRUCTION: Pick up the shoe/slipper which is placed in front of your feet.

GRADING: Please mark the lowest category which applies.

( )	( )	( )	( )	( )	
4	3	2	1	0	( )
able to pick up slipper safely and easily	able to pick up slipper but needs supervision	unable to pick up but reaches 1-2 inches from slipper and keeps balance indep	unable to pick up and needs supervision while trying	unable to try/ needs assist to keep from falling	

WHILE STANDING UNSUPPORTED

10 TURNING TO LOOK BEHIND OVER LEFT AND RIGHT SHOULDERS

INSTRUCTION: Turn to look behind you over toward left shoulder. Repeat to the right.

GRADING: Please mark the lowest category which applies.

( )	( )	( )	( )	( )	
4	3	2	1	0	( )
looks behind from both sides and weight shifts well	looks behind one side only other side shows less weight shift	turns sideways only but maintains balance	needs supervision when turning	needs assist to keep from falling	

11 TURN 360 DEGREES

INSTRUCTION: Turn completely around in a full circle. Pause. Then turn a full circle in the other direction.

GRADING: Please mark the lowest category which applies.

( )	( )	( )	( )	( )	
4	3	2	1	0	( )
able to turn 360 safely in < 4 sec each side	able to turn 360 safely one side only < 4 sec	able to turn 360 safely but slowly	needs close supervision or verbal cuing	needs assistance while turning	

12. DYNAMIC WEIGHT SHIFTING WHILE STANDING UNSUPPORTED

NUMBER OF TIMES ALTERNATE FOOT TOUCHES STOOL (7-8" high

INSTRUCTION: Place each foot alternately on the stool. Continue until each foot has touched the stool four times.

GRADING: Please mark the lowest category which ap

( )	( )	( )	( )	( )	
4	3	2	1	0	( )
able to stand indep and safely and complete 8 steps in 20 sec	able to stand indep and complete 8 steps > 20 sec	able to complete 4 steps without aid with supervision	able to complete >2 steps needs minimal assist	needs assistance to keep from falling/ unable to try	

13 STANDING UNSUPPORTED ONE FOOT IN FRONT

INSTRUCTION: (DEMONSTRATE to subject)

Place one foot directly in front of the other. If you feel that you cannot place your foot directly in front, try to step far enough ahead that the heel of your forward foot is ahead of the toes of the other foot.

GRADING: Please mark the lowest category which applies.

( )	( )	( )	( )	( )	
4	3	2	1	0	( )
able to place foot tandem indep and hold 30 sec	able to place foot ahead of other indep and hold 30 sec	able to take small step indep and hold 30 sec	needs help to step but can hold 15 sec	loses balance while stepping or standing	

14 STANDING ON ONE LEG

INSTRUCTION: Stand on one leg as long as you can without holding.

GRADING: Please mark the lowest category which applies.

( )	( )	( )	( )	( )	
4	3	2	1	0	( )
able to lift leg indep and hold >10 sec	able to lift leg indep and hold 5-10 sec	able to lift leg indep and hold =or>3 sec	tries to lift leg unable to hold 3 sec but remains standing indep	unable to try or needs assist to prevent fall	

TOTAL SCORE ( )

maximum = 56

TOTAL ___

APPENDIX 2

Contents:

BARTHEL SCALE

Patient Name: ______________________ Study #: ______________________

Location: ______________________ Room: ______________________

Date: ______________________ Examiner: ______________________

Evaluation: Initial _____; 2 _____; 3 _____; 4 _____

SELF CARE	Self	Some Aid	Can't Do	Score	Sub-total
1. Drinks from cup	4	0	0		
2. Eating	6	3	0		
3. Dress upper body	5	3	0		
4. Dress lower body	7	4	0		A.D.L.
5. Put on brace	0	-2	0		
6. Grooming	5	0	0		
7. Washing	6	0	0		
CONTINENCE					CONTINENCE
8. Bladder control	10	5	0		
9. Bowel control	10	5	0		
MOBILITY					
10. Transfer chair	15	7	0		MOBILITY
11. Transfer toilet	6	3	0		
12. Tub or shower	1	0	0		
13. Walks 50 yds.	15	10	0		
14. Stairs	10	5	0		
15. Wheeling if not walking	5	0	0		

TOTAL (100) ______

Guide to scoring BARTHEL INDEX

Self Care

1. Drinks from cup

Ind	4	Drinks from a cup; pours liquids and opens a milk carton
Help	0	Drinks and swallows but requires the presence or help from someone else for supervision, cueing, coaxing or physical assistance during the activity.
Null	0	Cannot drink but must rely at least in part on fluid intake through other routes such as parenteral or gastrostomy feedings.

2. Eating

Ind	6	Eats from a dish on a tray or table as customarily prepared and served; cuts meat and butters bread. May use adaptive or assistive devices.
Help	3	Takes full meals by mouth (ie chews and swallows) but requires the presence or help from someone else for supervision, cueing, coaxing or physical assistance during the activity. Does not rely on alimentation through other routes such as parenteral or gastrostomy feedings.
Null	0	Cannot take full meals by mouth but must rely on alimentation through other routes such as parenteral or gastrostomy feedings.

3. Dress upper body

Ind	5	Dresses and undresses upper body; able to handle bra, slip, pull-over garment and front-opening garment; able to manage zippers, buttons and snaps. Manages in reasonable time.
Help	3	Performs at least half the effort himself but needs help in putting on and removing or fastening.
Null	0	Does not assist or dressing is not performed.

4. Dress lower body

Ind	7	Dresses and undresses lower body, able to handle underpants, slacks, skirt, belt, stockings, and shoes, able to manage zippers, buttons, snaps, and garters. May use special closure.
Help	4	Performs at least half the effort himself, but needs help in putting on or removing or fastening.
Null	0	Does not assist or dressing is not performed.

5. Put on brace

Ind	0	Dons prescribed sling, splint, brace, orthosis, corset or prosthesis with reasonable time.
Help	-2	Needs partial or complete assistance to don brace, orthosis or prosthesis.
Null	0	Does not have brace, orthosis or prosthesis.

6. Grooming

Ind	5	Cleans teeth or dentures, combs and brushes hair, shaves, applies make-up, including all preparations.
Help	0	Assistance, supervision or cueing given for activity.
Null	0	Does not perform grooming activities; complete assistance given.

7. Wash

Ind	5	Washes and dries face and entire body (except shampooing hair), including taking water if bath is taken away from the sink or tub. May use adaptive or assistive devices.
Help	0	Assistance, supervision, cueing, coaxing given in washing and drying and inspecting skin of body.
Null	0	Does not wash and dry and inspect skin of body, or complete assistance is given.

Continence

8. Bladder continence

Ind	10	Has control of bladder (no accidents)
Help	5	Needs assistance with external device, or has occasional accidents; or cannot wait to get to bed pan or the toilet in time; another person is required to help in maintaining a voiding or excretion pattern or else the individual has occasional sphincter accidents but not on a daily basis.
Null	0	Is frequently wet due to incontinence despite aids or assistance. Despite aids or help from another person. The individual is wet on a frequent or almost daily basis. This might necessitate the wearing of diapers or other absorbent pads. This does not include the simple presence of a catheter or ostomy device provided that it is well maintained such that wetting of skin, clothing or bedding does not occur.

9.Bowel continence

Ind	10	Has control of bowels (no accidents)
Help	5	Needs assistance using suppository or taking an enema or has occasional accidents; another person is required to help in maintaining an excretion pattern or else the individual has occasional sphincter accidents but not on a daily basis.
Null	0	Is frequently soiled due to incontinence despite aids or assistance. Despite aids or help from another person, the individual is soiled on a frequent or almost daily basis. This might necessitate the wearing of diapers or other absorbent pads. This does not include the simple presence of an ostomy device provided that it is well maintained such that soiling of skin, clothing or bedding does not occur.

Mobility

10. Transfer chair

Ind	15	Approaches, sits down or gets up from a regular chair safely; if in wheelchair, able to approach a bed or another chair, lock brakes, lift foot rest, and safely perform either a standing pivot or sliding transfer; able to return safely, changing the position of the wheel chair if necessary; able to remove and replace arm rest if necessary.
Help	7	Minimal lifting required and/or supervision, cueing, coaxing.
Null	0	Heavy lifting is required; complete assistance is given.

11. Transfer toilet

Ind	6	Performs as in transferring to and from a chair with safety to toilet with fixed plumbing. May use an assistive device.
Help	3	Needs assistance for balancing as in adjusting clothes, minimal assistance or lifting is required.
Null	0	Heavy lifting is required, complete assistance is given.

12. Tub or bath

Ind	1	Enters and leaves a tub or shower tall independently and safely. May use assistive devices.
Help	0	Assistance, supervision, cueing given; minimal lifting involved.
Null	0	Heavy lifting required, complete assistnce given.

13. Walks 50 yards

Ind	15	Walks at least 50 yards; may use brace, cane, crutches, walker or special shoes.
Help	10	Walks at least 50 yards with assistance or supervision for safety.
Null	0	Does not walk 50 yards

14. Stairs

Ind	10	Goes up and down at least one flight of stairs without using any type of support and with safety. May wear brace.
Help	5	Goes up and down at least one flight of stairs but needs assistance or supervision for safety.
Null	0	Does not go up one flight of stairs; or is carried.

15. Wheeling if not walking

Ind	5	Propels wheelchair independently for at least 50 yards: must be able to go around corners, turn around, maneuver on rugs, over doorsills, etc
Help	0	Needs assistance in propelling or maneuvring wheelchair.
Null	0	Wheelchair is not ordinarily used for locomotion.

key: Ind - Independent
Help - Helper
Null - Null or unable to perform

FUGL-MEYER ASSESSMENT

Patient Name ______________

UPPER EXTREMITY

1. Reflex Activity -Flexors ___/2
   -Extensors ___/2

2. <u>Flexor Synergy</u>

Shoulder:	Retraction	___/2
	Elevation	___/2
	Abduction	___/2
	Ext. Rotation	___/2
	Elbow Flexion	___/2
	Forearm Supination	___/2

<u>Extensor synergy</u>

Shoulder Add., Int. Rot.	___/2
Elbow Extension	___/2
Forearm Pronation	___/2

3. Hand to lumbar spine	___/2
Shoulder Flexion to 90°	___/2
Pron/Supination, Elbow @ 90°	___/2
4. Shoulder Abduction to 90°	___/2
Shoulder Flexion from 90° to 180°	___/2
Pron/Supination, Elbow @ 0°	___/2
5. Normal Reflexes: biceps, triceps Finger flexors	___/2

Elbow 0° - wrist stable	___/2
Elbow 0° - wrist flex/ext	___/2
Elbow 90° - wrist stable	___/2
Elbow 90° - wrist flex/ext	___/2
Circumduction	___/2

Mass finger flexion	___/2
Mass finger extension	___/2
MCP extension - Distal flexion	___/2
Thumb adduction	___/2
Thumb opposition	___/2
Cylindrical grasp	___/2
Spherical grasp	___/2

Tremor	___/2
Dysmetria	___/2
Speed	___/2

UPPER EXTREMITY TOTAL ___/66

LOWER EXTREMITY

1. Reflex Activity:	Knee	___/2
	Achilles	___/2

2. <u>Flexor Synergy</u>

Hip Flexion (Abd., Ext. Rot.)	___/2
Knee Flexion	___/2
Ankle Dorsiflexion	___/2

<u>Extensor Synergy</u>

Hip Extension	___/2
Hip Adduction	___/2
Knee Extension	___/2
Ankle Plantarflexion	___/2

3. Knee flexion > 90°	___/2
Ankle Dorsiflexion	___/2
4. Knee Flexion, Hip @ 90°	___/2
Ankle Dorsiflexion	___/2
5. Normal Reflexes: Knee flexors, extensors Achilles	___/2

Tremor	___/2
Dysmetria	___/2
Speed	___/2

LOWER EXTREMITY TOTAL ___/34

BALANCE

Sitting unsupported	___/2
Protective reaction: less affected side	___/2
Protective reaction: more affected side	___/2
Supported standing	___/2
Unsupported standing	___/2
Standing: less affected leg	___/2
Standing: more affected leg	___/2

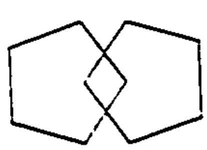
BALANCE TOTAL ___/14

The SENSATION and PASSIVE R.O.M. scales of the original Fugl-Meyer assessment have not been reprinted here. They have been shown to be too unreliable for useful clinical measures, and they are infrequently used.

NAME: DATE: STUDY NUMBER:

MINI-MENTAL STATE EXAMINATION

			SCORE	POINTS
Orientation				
1.	What is the	Year?	______	1
		Season?	______	1
		Date?	______	1
		Day?	______	1
		Month?	______	1
2.	Where are we?	Country?	______	1
		Province?	______	1
		City?	______	1
		Building?	______	1
		Floor?	______	1
Registration				
3.	Name three objects, taking one second to say each. Then ask the patient all three after you have said them. Give one point for each correct answer. Repeat the answers until the patient learns all three.		______	3
Attention and Calculation				
4.	Serial sevens. Give one point for each correct answer. Stop after five answers. Alternate spell WORLD backwards.		______	5
Recall				
5.	Ask for names of three objects learned in Question 3. Give one point for each correct answer.		______	3
Language				
6.	Point to a pencil and a watch. Have the patient name them as you point.		______	2
7.	Have the patient repeat "No ifs, ands or buts."		______	1
8.	Have the patient follow a 3-stage command: "Take a paper in your right hand. Fold the paper in half. Put the paper on the floor."		______	3
9.	Have the patient read and obey the following: CLOSE YOUR EYES (Write in large letters)		______	1
10.	Have the patient write a sentence of his or her choice. (The sentence should contain a subject and an object, and should make sense. Ignore spelling errors when scoring.		______	1
11.	Enlarge the design printed below to 1.5 cm per side, and have the patient copy it. Give one point if all sides and angles are preserved and if the intersecting sides form a quadrangle.)		______	1
		TOTAL	______	30



APPENDIX 3

Contents:

Note: Separate consent and screening forms were used for the Montreal General and the Royal Victoria/MNI but only one sample is included in the Appendix.

School of Physical and Occupational Therapy

McGill University

Griffith-McConnell Residence

I ______________________________ agree to participate in this research study on balance.

The purpose of this study is to test a scale that measures a person's balance ability. This scale will judge balance by asking you to perform movements common in everyday life. Your participation in the study will help us determine if this is a good scale to measure the ability of older people to maintain their balance while performing normal activities.

I understand that I will be asked to perform movements while sitting and standing according to my level of ability. These movements include turning, reaching forward and changing foot positions. In addition, I will be asked about my ability to do self-care activities such as dressing. The length of the interview will be approximately 15 minutes on one day, and 15 the following day, depending on each person's ability and the number of rests required. I understand that I may rest at any time and that I may end the interview at any time. My refusal to participate or to continue the testing will not affect any other treatments I receive. I understand that the measurements will be repeated on three occasions over a nine month period to check for changes in balancing scores. These visits will be made wherever I am residing at that point in time.

All information will be kept confidential and the results known only to the investigators. I have been told that this study has been approved by the Ethics Review Committees of McGill University and the Griffith-McConnell Residence. The risk involved is no greater than the risk associated with my activities of daily living. The possible benefit of my participation is that the scale will eventually be useful for other elderly persons. I understand that any enquiries I may have about this study will be answered. I can direct my questions to:

Ms Kathy Berg or Dr. Sharon Wood-Dauphinee
phone: 398-4500 or 398-4523
School of Physical and Occupational Therapy
McGill University
3654 Drummond
Montreal, Quebec

Date:____________________ Signature:______________________________

As research assistant, I declare that I have fully explained the objectives of the study and the extent of involvement requested of the resident.

____

SCHOOL OF PHYSICAL AND OCCUPATIONAL THERAPY
McGILL UNIVERSITY
GRIFFITH-MCCONNELL RESIDENCE

Researchers at the School of Physical and Occupational Therapy at McGill University in collaboration with the Griffith McConnell Residence are conducting a study designed to test a scale that measures a person's ability to balance.

Because your relative is at this time unable to fully understand the purpose of the study, we are asking you to sign this awareness which essentially states that you are aware of the study and of your relative's participation in it.

In this study, your relative will be asked to perform movements while sitting and standing according to his/her level of ability. These movements include turning, reaching forward and changing foot positions. In addition, he/she will be asked to do simple exercises with his/her arm and leg, and discuss his/her ability to do self-care activities such as dressing. The length of the test will be approximately 15 minutes, depending on each person's ability. Your relative may rest at any time, and he/she may end the interview at any time, Your relative's refusal to participate or to continue the testing will not affect any other treatments he/she receives, The measurements will be repeated on three occasions over a 1-year period to check for changes in balance scores.

All information obtained from the resident or from the charts will be kept confidential and the results known only to the investigators. This study has been approved by the Ethics Review Committees of McGill University and the Griffith McConnell Residence. The risk involved is no greater than the risk associated with your relative's activities of daily living. The possible benefit of his/her participation is that the scale will eventually be useful for other patients.

Date:______________ Signature:___________________________

Witness:______________________

Enquiries about this study can be directed to:

Katherine Berg or Dr Sharon Wood-Dauphinee

Phone: 398-4500

School of Physical and Occupational Therapy

McGill University

3654 Drummond

Montreal, Quebec

H3G 1Y5

School of Physical and Occupational Therapy
McGill University
Montreal General Hospital

I ______________________________ agree to participate in this research study on balance.

The purpose of this study is to test a scale that measures a person's balance ability. This scale will judge balance by asking you to perform movements common in everyday life. Your participation in the study will help us determine if this is a good scale to measure the ability of older people to maintain their balance while performing normal activities.

I understand that I will be asked to perform movements while sitting and standing according to my level of ability. These movements include turning, reaching forward and changing foot positions. In addition, I will be asked to do simple exercises with my arm and leg, and discuss my ability to do self-care activities such as dressing. The length of the test will be approximately 30 minutes on one day, and 15 the following day, depending on each person's ability and the number of rests required. I understand that I may rest at any time and that I may end the interview at any time. My refusal to participate or to continue the testing will not affect any other treatments I receive. I understand that the measurements will be repeated on three occasions over a three month period to check for changes in balancing scores. These visits will be made wherever I am residing at that point in time.

All information obtained from the patients or from the charts will be kept confidential and the results known only to the investigators. I have been told that this study has been approved by the Ethics Review Committees of McGill University and the Montreal General Hospital. The risk involved is no greater than the risk associated with my activities of daily living. The possible benefit of my participation is that the scale will eventually be useful for other patients. I understand that any enquiries I may have about this study will be answered. I can direct my questions to:

Ms Katherine Berg or Dr. Sharon Wood-Dauphinee

phone: 398-4500 or 398-4523
School of Physical and Occupational Therapy
McGill University
3654 Drummond
Montreal, Quebec

or Mrs Jacqueline Harvey
Physiotherapy Department : 937-6011 local 2900
Montreal General Hospital

Date:________________ Signature:__________________________

Witness:__________________________

Ecole de physiothérapie et d'ergothérapie
Université McGill
Hôpital Général de Montréal

Je, sousigné(e), ______________________________ accepte de participer à l'étude sur l'équilibre.

Cette étude vise à évaluer une échelle de mesure de l'aptitude d'une personne à l'équilibre. Cette échelle permettra d'évaluer l'équilibre chez des personnes à qui on demande d'effectuer les mouvements qu'elles exécutent dans la vie courante. Ma participation à l'étude aidera à déterminer la valeur de cette échelle pour mésurer l'aptitude de personnes agées à maintenir l'équilibre lors d'activités de la vie quotidienne.

Je comprends qu'on me demandera d'effectuer des mouvements, en position assise et en position debout, selon mon niveau d'aptitude. Ces mouvements comprendront: rotation, étirement et changement de position des pieds. De plus, on me demandera d'exécuter des exercises simples, des jambes et et des bras, et on discutera avec moi de mon aptitude à prendre en charge mes soins personnels, tel m'habiller. L'entrevue durera environ 30 minutes la première journée et 15 minutes la suivante, selon l'aptitude de chacune des personnes. On m'a affirmé que je pourrais me reposer ou mettre fin à l'entrevue à tout moment. Mon refus de participer ou de poursuivre l'évaluation n'affectera en rien les autres traitements. Les mesures pourront être répétées subséquemment à trois reprises au cours des trois prochains mois, afin de vérifier si des changements surviennent dans l'aptitude à l'équilibre.

Tous les renseignements personnels obtenus du patient ou de son dossier demeureront confidentiels. On m'a affirmé que cette étude avait reçu l'approbation du comité d'éthique de l'université McGill et l'hôpital Général de Montréal. Les risques que comporte cette étude sont du même ordre que les risques que je cours quotidiennement. L'avantage possible de ma participation est que cette échelle servira éventuellement à d'autres malades et permettra aux professionnels de la santé d'évaluer objectivement diverses techniques de traitement. Il est entendu qu'on répondra à toutes les questions que j'aurai au sujet de l'étude, que je peux adresser directement au

Dr. Sharon Wood-Dauphinée ou à Mlle Katherine Berg:
Ecole de physiothérapie et d'ergothérapie: 398-4500 ou 398-4514
Université McGill, 3654 rue Drummond, Montréal, PQ H3G 1Y5
ou
Mme Jacqueline Harvey
Département de physiothérapie :937-6011 poste 2900
Hôpital Général de Montréal

Date: Signature:____________________

Témoin: ________________________

SCHOOL OF PHYSICAL AND OCCUPATIONAL THERAPY

McGILL UNIVERSITY

ROYAL VICTORIA HOSPITAL and the MONTREAL NEUROLOGICAL HOSPITAL

Researchers at the School of Physical and Occupational Therapy at McGill in collaboration with the Physiotherapy Department of the RVH and MNH are conducting a study designed to test a scale that measures a person's balance ability.

Because your relative is at this time unable to fully understand the purpose of the study, we are asking you to sign this awareness form which essentially states that you are aware of the study and of your relative's participation in it.

If he/she participates in this study your relative will be asked to perform movements while sitting and standing according to his/her ability. These movements include turning, reaching forward and changing foot positions. In addition, he/she will be asked to do simple exercices with his/her arm and leg, and discuss his/her ability to do self-care activities such as dressing. The length of the test will be approximately 30 minutes on one day, and 15 minutes the following day, depending on each person's ability and the number of rests required. Your relative may rest at any time, and he/she may end the interview at any time. Your relative's refusal to participate or to continue the testing will not affect any other treatments he/she receives. The measurements will be repeated on three occasions over a three-month period to check for changes in balancing scores. These visits will be made wherever your relative is residing at that point in time.

All information obtained from the patient or from the charts will be kept confidential and the results known only to the investigators. This study has been approved by the Ethics Review Committees of McGill University, the Royal Victoria Hospital and the Montreal Neurological Hospital. The risk involved is no greater than the risk associated with your relative's activities of daily living. The possible benefit of his/her participation is that the scale will eventually be useful for other patients. Enquiries about this study can be directed to:

Ms Kathy Berg or Dr. Sharon Wood-Dauphinee

phone: 398-4500 or 398-4523

School of Physical and Occupational Therapy

McGill University

3654 Drummond

Montreal, Quebec

or Dr. David Gayton

phone: 842-1231, local 4677

Royal Victoria Hospital

Date:________________ Signature ____________________________

Witness: ______________________

UNIVERSITÉ McGILL

École de Physiothérapie et d'Ergothérapie

L'Hôpital Général de Montréal

Des chercheurs de l'École de physiothérapie et d'ergothérapie de l'Université McGill en collaboration avec le département de physiothérapie de l'Hôpital Général de Montréal sont impliqués dans une étude qui vise à évaluer une échelle de mesure de l'aptitude d'une personne à l'équilibre.

Parce que Monsieur ou Madame n'est pas en mesure de comprendre le but de cette étude, nous vous demandons de signer cette formule. En apposant votre signature, vous aurez pris connaissance des buts de l'étude et vous serez conscient que Monsieur ou Madame y participe.

Si la personne participe à cette étude, il lui sera demandé d'effectuer des mouvements, en position assise et en position debout, selon son niveau d'aptitude. Ces mouvements comprendront: la rotation, l'étirement et un changement de position des pieds. De plus, il/elle devra exécuter des exercices simples, des jambes et des bras, et on discutera avec cette personne de son aptitude à prendre en charge ses soins personnels, tel s'habiller.

L'entrevue durera environ 30 minutes la première journée et 15 minutes la suivante, selon l'aptitude de chacune des personnes et les périodes de repos qui seront prises. La personne concernée pourra se reposer ou mettre fin à l'entrevue à tout moment.

Le refus de cette personne de participer ou de poursuivre l'évaluation n'affectera en rien ses autres traitements. Les mesures seront répétées subséquemment à trois reprises au cours des trois prochains mois, afin de vérifier si des changements surviennent dans l'aptitude à l'équilibre. Ces visites s'effectueront au lieu de résidence de la personne.

Tous les renseignements personnels obtenus du patient ou de son dossier demeuront confidentiels et les résultats seront dévoilés aux chercheurs seulement. Cette étude a reçu l'approbation du Comité d'éthique de l'Université McGill, et de l'Hôpital de Montréal. Les risques que comporte cette étude sont du même ordre que les risques qu'une personne court quotidiennement. L'avantage possible de la participation de cette personne est que cette échelle servira éventuellement à d'autres patients. Pour plus de renseignements sur cette étude, veuillez vous adresser directement à:

Mlle Katherine Berg ou Dr. Sharon Wood-Dauphinee au:
398-4500
École de physiothérapie
et d'ergothérapie
Université McGill
3654 rue Drummond
Montréal, PQ H3G 1Y5

Date: ______________________ Signature: ________________________

Témoin: ______________________

GRIFFITH MCCONNELL SCREENING

Screening number _ _ _

SURNAME: Forename:

TELEPHONE:

LOCATION: ROOM NUMBER:

SEX: AGE:

CRITERIA	YES	NO
AGE> = 60	☐	___
HAS DEGREE OF INDEPENDENT MOBILITY	☐	___
-usually walks without aids ___		
-uses cane when outdoors ___		
-uses cane indoors ___		
-uses walker usually ___		
-needs assist of a person ___ but can reach toilet unassisted with a W/C or mobility aids		
IS MEDICALLY STABLE	☐	___

IF ALL THREE BOXES ARE CHECKED, PATIENT IS ELIGIBLE FOR THE STUDY.

INFORMED CONSENT

Resident understands objectives of the study and has signed consent form _____ (yes, no)

If no:

Resident has a MiniMental Score of ______ ; family aware _____

Name of next of kin: Telephone:

IF NO CONSENT, REASON:

ADMIT TO STUDY: DATE:

RVH-MNI SCREENING

verify following information:

Patient name:
hospital number:
ward: room:
age: sex:
Physician:

SCREENING NUMBER: ___ ___ ___

DATE OF STROKE COMPLETION:

CRITERIA	YES	NO
AGE => 40	☐	___
0- 14 DAYS POST SYMPTOM ONSET	☐	___
ELIGIBLE FOR MEDICARE	☐	___
FOLLOW UP VISIT POSSIBLE	☐	___
(lives in Montreal, or will be here for 3 months)	☐	___
MEDICALLY STABLE - allowed to sit up as tolerated	☐	___
PRESENCE MEDICAL PROBLEMS THAT MAY INTERFERE WITH REHABILITATION - Alzheimer's, legal blindness, amputee functionally dependent before CVA, Barthel <40	☐	___
CLINICAL DIAGNOSIS STROKE	☐	___
REFERRAL SENT TO PT	☐	___

IF ALL 8 BOXES ARE CHECKED, PATIENT IS ELIGIBLE FOR THE STUDY.

INFORMED CONSENT

Patient understands the objectives of study and has signed consent form______
yes, no

If no, awareness form signed by next of kin: ________ (yes,no)

name of next of kin: telephone:

If no, reason:

ADMIT TO STUDY: DATE:

PROFILE SHEETS

STUDY NUMBER	___ ___ ___	1-3
Soc/clinical	1	4
STUDY SECTION: Yes = 1; No = 2		
Predictive	___	5
Construct	___	6
Reliability	___	7
Concurrent	___	8
Age(yrs)	___ ___	9-10
Sex male: 1; female: 2	___	11
Language: French: 1; Eng.: 2; Other: 3	___	12
Marital stat: Married: 1; Never married: 2; Formerly married: 3	___	13
Living arrangement: alone,1; family,2; residence 3; institution,4; other 5	___	14
Years of schooling	_/_	15-16
Employment: employed FT: 1; part-time: 2; retired: 3; unemployed: 4; housewife: 5; NA: 9	___	17
Occupation (main during lifetime) prof/manager: 1; clerical: 2; sales: 3; service/rec: 4; transportation/commun.: 5; crafts/production: 6; laborer: 7; student: 8; housewife: 9; none: 10	_/_	18-19
Diagnosis: Yes = 1, No = 2		
CVA	___	20
Parkinson	___	21
Other neurological	___	22
Cardiovascular	___	23
Pulmonary	___	24
Diabetes	___	25
Periph vasc disease (Amp)	___	26
Rheumatic disease	___	27
Visual problems	___	28
Hypertension	___	29
Renal disease	___	30
Gastrointestinal	___	31
Genitourinary	___	32
Neoplastic	___	33
Orthopaedic history (fractures, arthroplasty)	___	34
Alcoholism	___	35
Other	___	36

Meds: No = 0; if yes write # taken in category		
Major tranquilizer	___	37
Minor tranquilizer	___	38
Anti-platelet/Anti-coagulant	___	39
Anti-depressants	___	40
Anti-hypertensives	___	41
Ischemic heart disease	___	42
Anti-inflammatory	___	43
Other:	___	44
Completed study: yes 1, no 2;	___	45
If no, reason: Incapacitating [illegible] 1, incapacitating disease 2, voluntary withdrawal 3, died 4	___	46
Date death or premature D/c	__|__|__ (d m y)	47-52
Current stroke episode date symptom onset	__|__|__ (d m y)	53-58
Side of weakness R = 1; L = 2; Both = 3	___	59
Use mobility aids: usually uses no mobility aid (1) usually uses cane outdoors (2) usually uses cane indoors (3) usually uses quad cane or walker (4) requires assistant to walk (5) independent with wheelchair (6)	___	60
Mini-mental state	_/_	61-62

Study number ___ ___ ___ 1-3

Initial Evaluation Info _2_ 4

Location:
GMcC 1, RVH-MNI 2, home 3,
convalescent 4, other 5 ___ 5

Fall : yes 1, no 2

with injury	___	6
no injury	___	7

Total falls
in interval ___ ___ 8-9

Date balance
evaluation

__ __ | __ __ | __ __ 10-15
d m yr

Indep rater # ___ ___ 16-17

Balance I

sit to stand	1	___	18
stand	2	___	19
sit	3	___	20
stand to sit	4	___	21
transfer	5	___	22
stand EC	6	___	23
stand FT	7	___	24
arm reach	8	___	25
pick up	9	___	26
twist turn	10	___	27
turn 360	11	___	28
stool step	12	___	29
tandem	13	___	30
one leg	14	___	31

Date initial assessment
of Barthel

__ __ | __ __ | __ __ 32-37
d m yr

BARTHEL

Mobility	_/_	38-39
ADL	_/_	40-41
Cont	_/_	42-43

Global rating ___ 44

poor = 1
fair = 2
good = 3
not given = 4

Profession of caregiver ___ 45

Nurse = 1
PT = 2
OT = 3
Nursing asst = 4
N/A = 9

Self perception patient ___ 46

Poor = 1
Fair = 2
Good = 3
Not given = 9

Fugl-meyer I

Arm	___ ___	47-48
Leg	___ ___	49-50
Balance	___ ___	51-52

Study number ___ ___ ___ 1-3

Inter- Reliability study _6_ 4

Location:
GMcC 1, RVH-MNI 2, home 3,
convalescent 4, other 5 ___ 5

Fall : yes 1, no 2

with injury ___ 6
no injury ___ 7

Total falls
in interval ___ ___ 8-9

Date reliability balance
evaluation

_ _|_ _|_ _ 10-15
d m yr

Rater # ___ ___ 16-17

Caregiver Balance Scores

sit to stand	1	___	18
stand	2	___	19
sit	3	___	20
stand to sit	4	___	21
transfer	5	___	22
stand EC	6	___	23
stand FT	7	___	24
arm reach	8	___	25
pick up	9	___	26
twist turn	10	___	27
turn 360	11	___	28
stool step	12	___	29
tandem	13	___	30
one leg	14	___	31

Profession of caregiver ___ 32
Nurse = 1
PT = 2
OT = 3
Nursing asst = 4
N/A = 9

Date corresponding balance
evaluation

_ _|_ _|_ _ 33-38
d m yr

Rater # ___ ___ 39-40

Indep rater Balance Scores

sit to stand	1	___	41
stand	2	___	42
sit	3	___	43
stand to sit	4	___	44
transfer	5	___	45
stand EC	6	___	46
stand FT	7	___	47
arm reach	8	___	48
pick up	9	___	49
twist turn	10	___	50
turn 360	11	___	51
stool step	12	___	52
tandem	13	___	53
one leg	14	___	54

Study number ___ ___ ___ 1-3

Intra- Reliability study _7_ 4

Location:
GMcC 1, RVH-MNI 2, home 3,
convalescent 4, other 5 ___ 5

Fall : yes 1, no 2
with injury ___ 6
no injury ___ 7
Total falls
in interval ___ ___ 8-9

Date repeat balance
evaluation
_ _|_ _|_ _ 10-15
d m yr
Rater # ___ ___ 16-17

Repeat Balance Scores

sit to stand	1	___	18
stand	2	___	19
sit	3	___	20
stand to sit	4	___	21
transfer	5	___	22
stand EC	6	___	23
stand FT	7	___	24
arm reach	8	___	25
pick up	9	___	26
twist turn	10	___	27
turn 360	11	___	28
stool step	12	___	29
tandem	13	___	30
one leg	14	___	31

Date corresponding balance
evaluation
_ _|_ _|_ _ 32-37
d m yr
Rater # ___ ___ 38-39

Indep rater Balance Scores

sit to stand	1	___	40
stand	2	___	41
sit	3	___	42
stand to sit	4	___	43
transfer	5	___	44
stand EC	6	___	45
stand FT	7	___	46
arm reach	8	___	47
pick up	9	___	48
twist turn	10	___	49
turn 360	11	___	50
stool step	12	___	51
tandem	13	___	52
one leg	14	___	53

FALL INCIDENT REPORT

DATE OF FALL:

PLEASE DESCRIBE THE CIRCUMSTANCES OF THE FALL:

If they have not already given the following information, ask

What were you doing when you fell (eg going to the bathroom, hanging clothes)

Where were you? (eg outside shopping, in the dining room)

In own room ____
In hallway ____
In bathroom ____
Outdoors ____
other ____________________

What time of day was it? (record approximate time in 24 hour)

What kind of shoes or boots were you wearing at the time?

usual walking shoe ____ Height of heel of usual shoe ____
slipper ____
winter boots ____
bare feet ____

Were you using your usual walking aid? yes ___ no ___

Your usual mobility aid is: Cane ___
walker ___
quad cane ___
wheelchair ___
does not use an aid ____

Were you in a hurry, or were you startled or distracted?

Was there any obstacle or external disturbance

APPENDIX 4.1

Means and standard deviations (SD) of the Balance Scale Scores of the elderly residents within each category of global rating at each evaluation point

Global Rating	EVALUATION POINTS							
	N	At entry Mean (SD)	N	3 Months Mean (SD)	N	6 Months Mean (SD)	N	9 Months Mean (SD)
GOOD	63	50.2 (3.47)	55	49.4 (4.31)	55	49.8 (4.70)	53	49.7 (3.53)
FAIR	40	43.1 (7.08)	39	44.7 (6.80)	34	44.6 (6.21)	33	44.1 (6.64)
POOR	10	40.8 (7.50)	11	38.0 (9.85)	11	33.5 (8.24)	9	39.5 (11.39)
		rho = .61		rho = .50		rho = .62		rho = .52
		$F_{df2,110}$ 27.8 $p < .0001$		$F_{df2,103}$ 19.4 $p < .0001$		$F_{df2,94}$ 39.7 $p < .0001$		$F_{df2,92}$ 17.3 $p < .0001$

APPENDIX 4.2

Sociodemographic and functional characteristics of the residents at baseline grouped by their fall status at the end of the year (N=113)

	Non Fallers N=60 **Mean (SD)** or Number (percent)	Single Fallers N=24 **Mean (SD)** or Number (percent)	Multiple Fallers N=29 **Mean (SD)** or Number (percent)
Age	**82.8 (5.2)**	**83.3 (4.9)**	**85.1 (5.1)**
Sex			
Female	49 (81.7)	20 (83.3)	24 (82.8)
Mental Status MMSE	**28.1 (2.7)**	**27.9 (2.7)**	**27.5 (2.7)**
Barthel Mobility	**46.5 (1.5)**	**46.1 (2.5)**	**44.5 (4.2)***
Balance	**48.8 (4.4)**	**46.8 (7.6)**	**42.8 (7.7)***
History of falls in past three months	3 (5.0)	3 (12.5)	10 (34.5)**
Mean # Conditions	**3.8 (1.4)**	**3.4 (1.3)**	**4.3 (1.4)**
Neurological	26 (43.3)	10 (41.7)	10 (34.5)
CVD	36 (60.0)	8 (33.3)	19 (65.5)**
Pulmonary	9 (15.0)	5 (20.8)	2 (6.9)
Diabetes	3 (5.0)	1 (4.2)	5 (17.2)
PVD	8 (13.3)	2 (8.3)	1 (3.4)
Rheumatic	24 (40.0)	12 (50.0)	13 (44.8)
Visual	21 (35.0)	7 (28.2)	17 (58.6)**
Hypertension	31 (5.7)	11 (45.8)	17 (58.6)
Gastrointestinal	11 (18.3)	9 (37.5)	6 (20.7)
Genitourinary	4 (6.7)	2 (8.3)	3 (10.3)
Neoplasm	9 (15.0)	1 (4.2)	6 (20.7)
Orthopaedic	16 (26.7)	2 (8.3)	10 (34.5)
Other	30 (50.0)	12 (50.0)	16 (55.2)
Mean # Medications	**4.3 (1.8)**	**3.2 (1.8)**	**3.6 (2.2)**
Tranquilizers	4 (6.7)	0 (0)	2 (6.9)
Sedatives	25 (41.7)	11 (45.8)	6 (20.7)
Antiplatelet	17 (28.3)	5 (20.8)	6 (20.7)
Antidepressant	6 (10.0)	1 (4.2)	3 (10.3)
Antihypertensive	32 (53.3)	7 (29.2)	12 (41.4)
Cardiac	28 (46.7)	6 (25.0)	14 (48.3)
Antiinflammatory	20 (50.0)	3 (12.5)	13 (44.8)**
Other	40 (78.3)	18 (75.0)	20 (69.0)

* F_{ratio} $p < .01$
** x^2 $p < .05$

APPENDIX 4.3

Sociodemographic and medical characteristics of stroke patients who did and did not complete the study (N=70)

	Completed Study N=60		Did Not Complete Study N=10	
	Mean (SD) Number (percent)		**Mean (SD)** Number (percent)	
Age	**70.8**	**(9.8)**	**76.6**	**(10.5)**
Sex				
Female	29	(51.7)	5 (50.0)	
Side of Stroke				
Right	32	(53.3)	6	(60.0)
Left	28	(46.7)	4	(40.0)
Usual Living Arrangements				
Alone	18	(30.0)	3	(30.0)
Family	40	(66.7)	6	(60.0)
Residence	1	(1.7)	1	(10.0)
Institution	1	(1.7)	0	
Language				
French	19	(31.7)	3	(30.0)
English	22	(36.7)	6	(60.0)
Other	19	(31.7)	1	(10.0)
Education	**8.5**	**(3.6)**	**8.8**	**(2.6)**
Mean # Conditions	**2.7**	**(1.4)**	**2.1**	**(1.2)**
Neurological	9	(15.0)	2	(20.0)
Cardiovascular	35	(58.3)	5	(50.0)
Pulmonary	8	(13.3)	0	(0.0)
Diabetes	20	(33.3)	2	(20.0)
Peripheral vascular	7	(11.7)	1	(10.0)
Rheumatic	4	(6.7)	0	(0.0)
Visual	5	(8.3)	0	(0.0)
Hypertension	39	(65.0)	7	(70.0)
Renal	2	(3.3)	2	(20.0)
Gastrointestinal	10	(16.7)	1	(10.0)
Genitourinary	6	(10.0)	0	(0.0)
Orthopaedic	4	(6.7)	0	(0.0)
Other	9	(15.0)	0	(0.0)
Mean # Medications	**3.8**	**(1.8)**	**3.8**	**(2.4)**
Sedatives	7	(11.7)	2	(20.0)
Antiplatelet/Anticoagulant	45	(75.0)	9	(90.0)
Antidepressant	1	(1.7)	1	(10.0)
Antihypertensive	30	(50.0)	6	(60.0)
Cardiac	31	(51.7)	4	(40.0)
Antiinflammatory	4	(6.7)	1	(10.0)
Other	41	(68.3)	5	(50.0)

Sociodemographic and clinical characteristics of the elderly residents in the reliability study (N=31)

Sociodemographic Characteristics	Mean (SD) Number (percent)	Clinical Characteristics	Mean (SD) Number (percent)
Age	**84.4 (5.0)**	Medical Problems	
		Neurological	13 (41.9)
Sex		Cardiovascular	15 (48.4)
Male	8 (25.8)	Pulmonary	8 (25.8)
Female	23 (74.2)	Diabetes	4 (12.9)
		Peripheral Vascular	5 (16.1)
Marital Status		Rheumatic	13 (41.9)
Married	5 (16.1)	Visual	12 (38.7)
Never Married	8 (25.8)	Hypertension	18 (58.1)
Formerly Married	18 (58.1)	Gastrointestinal	5 (16.1)
		Genitourinary	2 (6.5)
Language		Neoplasm	2 (6.5)
French	1 (3.2)	Orthopaedic	3 (9.7)
English	30 (96.8)	Other	18 (58.1)
Usual Living Arrangements		Mean # Diagnoses	**3.8 (1.2)**
Residence	31 (100.0)	Mental Status (MMSE)	**27.5 (2.9)**
Usual Occupation		Mobility Aids	
Professional	16 (51.6)	None	17 (54.8)
Clerical	10 (32.3)	Cane outdoors	7 (22.6)
Sales	1 (3.2)	Cane indoors	7 (22.6)
Transportation	1 (3.2)		
Production	1 (3.2)	Medications	
Labourer	1 (3.2)	Major tranquilizers	2 (6.5)
Housewife	4 (3.2)	Sedatives	10 (32.3)
		Antiplatelet/coagulant	7 (22.6)
Education (years)	**12.5 (3.3)**	Antidepressant	3 (9.7)
		Antihypertensive	16 (51.6)
		Cardiac	12 (38.7)
		Antiinflammatory	9 (29.0)
		Other	24 (77.4)
		Mean # Medications	**4.0 (2.2)**

Sociodemographic and clinical characteristics of the stroke patients in the reliability study (N=36)

Sociodemographic Characteristics	**Mean (SD)** Number (Percent)	Clinical Characteristics	**Mean (SD)** Number (Percent)
Age (years)	**72.4 (9.1)**	Side of stroke	
		Right	15 (41.7)
Sex		Left	21 (58.3)
Male	18 (50.0)		
Female	18 (50.0)	Sex	
		Male	18 (50.0)
Marital Status		Female	18 (50.0)
Married	20 (55.6)		
Never Married	4 (11.1)	Comorbidity	
Formerly Married	12 (33.3)	Neurological	8 (22.2)
		Cardiovascular	19 (52.8)
Language		Pulmonary	5 (13.9)
French	12 (33.3)	Diabetes	10 (27.8)
English	15 (41.7)	Peripheral Vascular	7 (19.4)
Other	9 (25.0)	Rheumatic	4 (11.1)
		Visual	5 (13.9)
Usual Living		Hypertension	23 (63.9)
Arrangements		Renal	1 (2.8)
Home Alone	13 (36.1)	Gastrointestinal	7 (19.4)
Family	20 (55.6)	Genitourinary	8 (22.2)
Residence	2 (5.6)	Orthopaedic	1 (2.8)
Institution	1 (2.8)	Other	5 (13.9)
Employment Status		Mean # Co-morbid	**2.9 (1.6))**
Full-time	1 (2.8)	Conditions	
Retired	30 (83.3)		
Housewife	5 (13.9)	Medications	
		Major tranquilizers	1 (2.8)
Usual Occupation		Sedatives	6 (16.7)
Professional	8 (22.9)	Antiplatelet/coagulant	25 (55.6)
Clerical	3 (8.6)	Antidepressant	2 (5.6)
Sales	2 (5.7)	Antihypertensive	17 (47.2)
Service	1 (2.9)	Cardiac	19 (52.8)
Transportation	2 (5.7)	Antiinflammatory	5 (13.9)
Production	5 (14.3)	Other	25 (69.4)
Labourer	4 (11.4)		
Housewife	9 (25.7)	Mean # Medications	**3.9 (2.0)**
None	1 (2.8)		
Education (years)	**10.1 (3.3)**		

APPENDIX 4.6

Corrected Item-to-Total Correlations and Internal consistency estimates for elderly residents across the evaluation points

Scale Item	EVALUATION POINTS			
	Baseline N=113	3 Months N=106	6 Months N=101	9 Months N=99
Sit to Stand	.641	.484	.564	.673
Standing	.405	.560	.604	.729
Stand to Sit	.630	.488	.640	.719
Transfer	.600	.610	.595	.749
Stand Eyes Closed	.555	.490	.452	.704
Stand Feet Together	.398	.572	.618	.522
Arm Reaching	.460	.527	.606	.629
Object Pick Up	.525	.468	.564	.729
Twisting	.624	.412	.506	.630
Turn 360°	.602	.633	.705	.683
Stepping	.592	.620	.623	.591
Tandem Standing	.378	.426	.501	.238
One Leg Standing	.437	.363	.438	.389
Cronbach's Alpha	.830	.839	.860	.836
Standardized Item Alpha	.871	.859	.889	.907

* One item was removed from this analysis. Sitting balance was scored as normal in all residents; therefore could not be included.

Corrected Item-to-Total Correlations and Internal consistency estimates for stroke patients across the evaluation points

Scale Item	EVALUATION POINTS			
	Baseline N=69	4 Weeks N=66	6 Weeks N=64	12 Weeks N=59
Sit to Stand	.894	.940	.882	.931
Standing	.933	.950	.928	.926
Sitting	.674	.652	.626	.645
Stand to Sit	.911	.937	.929	.903
Transfer	.923	.932	.918	.927
Stand Eyes Closed	.894	.935	.900	.922
Stand Feet Together	.886	.901	.919	.924
Arm Reaching	.923	.942	.926	.888
Object Pick Up	.884	.944	.945	.951
Twisting	.949	.932	.910	.910
Turn 360°	.859	.845	.878	.896
Stepping	.741	.771	.801	.792
Tandem Standing	.800	.854	.801	.813
One Leg Standing	.703	.748	.754	.765
Cronbach's Alpha	.973	.981	.979	.980
Standardized Item Alpha	.977	.981	.978	.980

Means and standard deviations of the item scores of the elderly residents at each evaluation

Scale Item	Baseline Mean (SD)	Three Months Mean (SD)	Six Months Mean (SD)	Nine Months Mean (SD)
Sit to Stand	3.56 (0.65)	3.66 (0.53)	3.64 (0.61)	3.66 (0.66)
Standing	3.91 (0.37)	3.84 (0.66)	3.93 (0.29)	3.87 (0.57)
Sitting	4.00 (0.00)	4.00 (0.00)	4.00 (0.00)	4.00 (0.00)
Stand to Sit	3.73 (0.55)	3.69 (0.67)	3.66 (0.59)	3.71 (0.52)
Transfer	3.75 (0.49)	3.76 (0.54)	3.70 (0.59)	3.73 (0.44)
Stand Eyes Closed	3.94 (0.24)	3.81 (0.69)	3.85 (0.55)	3.86 (0.57)
Stand Feet Together	3.67 (0.75)	3.66 (0.85)	3.51 (1.11)	3.66 (0.86)
Arm Reaching	2.90 (0.88)	3.08 (0.94)	3.20 (0.86)	3.43 (0.67)
Object Pick Up	3.75 (0.85)	3.72 (0.86)	3.63 (1.01)	3.73 (0.87)
Twisting	3.59 (0.84)	3.47 (0.97)	3.63 (0.78)	3.47 (0.92)
Turn 360°	3.46 (1.04)	3.42 (1.09)	3.29 (1.18)	3.34 (1.11)
Step on stool	2.62 (1.49)	2.63 (1.40)	2.51 (1.59)	2.71 (1.42)
Tandem Standing	2.29 (1.22)	2.30 (1.21)	2.33 (1.20)	2.21 (1.19)
One Leg Standing	1.66 (1.13)	1.46 (1.13)	1.42 (1.17)	1.45 (1.04)
	N=113	N=106	N=101	N=98