

VALIDITY OF PRE-DISCHARGE MEASURES FOR PREDICTING HARM

VALIDITY OF PRE-DISCHARGE FUNCTIONAL MEASURES FOR PREDICTING HARM
IN OLDER ADULTS

By ALISON DOUGLAS B.Sc.H., B.HSc.(OT), M.Sc.(OT)

A Thesis Submitted to the School of Graduate Studies
in Partial Fulfillment of
the Requirements for
the Degree Doctor of Philosophy, Rehabilitation Science

DOCTOR OF PHILOSOPHY (2011)
(Rehabilitation Science)

McMaster University
Hamilton, Ontario

TITLE: Validity of pre-discharge functional measures for predicting harm in older adults

AUTHOR: Alison Douglas B.ScH. (Queen's University), B.HSc. (O.T.) (McMaster University), M.Sc. (O.T.) (University of Alberta).

SUPERVISOR: Dr. Lori Letts PhD.

NUMBER OF PAGES: xi, 138

ABSTRACT

Older adults admitted to hospital often must make discharge decisions about whether they will be able to safely manage at home. Decision-making for clients and teams is supported by occupational therapy assessment. This thesis presents three manuscripts from a single study that was designed to address a need for evidence of validity of two measures for predicting harm.

The design was a prospective observational study in which older adults from an inpatient unit (n=47) were followed for six months for reports of incidents of harm. Baseline data included independent variables (e.g. age, sex, education, living alone, comorbidities, caregiving hours, ADL score) and scores on two measures: the Assessment of Motor and Process Skills (AMPS) and the Cognitive Performance Test (CPT).

The first manuscript contributed needed validation evidence for the CPT. The CPT correlated moderately with cognitive measures, and scores were not affected by age, and years of education. The CPT differentiated impaired persons differently from other measures. Results highlighted that further evidence to test the CPT against a criterion related to outcomes in the community was needed.

The second manuscript tested the trustworthiness of the outcome “incidents of harm”. Test-retest reliability was high and validation against daily logs and medical charts supported this method of measurement of incident of harm.

The third manuscript determined whether the AMPS and CPT were valid for predicting incidents of harm after discharge. The results showed that, compared with all independent variables, AMPS-Process scores were the most significant predictor of harm outcome. The CPT had a high specificity for identifying persons who did not have harm. Living alone, age and sex contributed to the prediction of harm. The implications of these results are that scores on the measures can inform patients, families and the team about older adults’ risk of incidents of harm after discharge.

ACKNOWLEDGEMENTS

I am extremely grateful for having had the support of my supervisor Lori Letts. She was able to understand when I needed support most, help negotiate the process of team research work and provide practical solutions. I am grateful for having had her guidance and encouragement throughout the process. Thank you also to Julie Richardson and Kevin Eva who were excellent mentors through their expertise, thoroughness, and willingness to be available for guidance.

It was wonderful to have the love and support of Jamie and James who were patient and encouraging. Thank you! Your consistent optimism made the journey enjoyable and your sense of humour allowed a balanced perspective on life.

Old friends Sharlene, Margie and Mary Ellen helped me to be committed and keep perspective on my own unique contribution to a profession and to others.

This thesis is dedicated to my parents who fostered my curiosity and love of learning.

PREFACE

The student's contribution to each work is summarized as follows:

For all three studies: Alison Douglas defined the research questions and provided the overall study design. She established a relationship with the research site, drafted the ethics submission, hired and managed the two research assistants and research grant budget, completed the data collection with the research assistants, analyzed the data and drafted the manuscripts.

Dr. Lori Letts assisted with editing and revising the ethics board submission, defining the allocation of data to the three manuscripts, interpreting the results and conceptual flow of the manuscripts.

Dr. Kevin Eva and Dr. Julie Richardson assisted with refining the objectives and protocols for each study, statistical analysis, interpretation of the findings and editorial assistance with writing the manuscripts.

TABLE OF CONTENTS

ABSTRACT	iii
ACKNOWLEDGEMENTS	iv
PREFACE	v
TABLE OF CONTENTS	vi
LIST OF TABLES.....	ix
LIST OF FIGURES.....	x
LIST OF APPENDICES.....	xi
CHAPTER 1: INTRODUCTION	1
Research Problem.....	1
Clinical Context and Background	1
Occupational therapy role with persons with dementia.....	2
Occupational therapists' use of cognitive assessment measures:.....	4
Types of evidence to support a measure	7
Study design and measure selection.....	11
Study Purpose and Objectives:.....	16
Summary:.....	22
References	23
CHAPTER 2: Use of the Cognitive Performance Test (CPT) for identifying deficits in hospitalized older adults.....	25
Abstract	26
Introduction.....	26

Methods.....	34
Results.....	39
Discussion.....	43
References.....	52
CHAPTER 3: Measurement of harm outcomes in older adults after hospital	
discharge: Reliability and Validity.....	55
Abstract.....	56
Introduction.....	57
Methods.....	60
Results.....	65
Discussion.....	69
Conclusions.....	72
References.....	74
CHAPTER 4: Validity of pre-discharge functional measures for predicting harm in	
older adults.....	76
Abstract.....	77
Introduction.....	78
Methods.....	82
Results.....	87
Discussion.....	99
Conclusion.....	106
CHAPTER 5: Discussion.....	114

Contributions of each manuscript	115
Overall Implications	119
Conclusion	134
References	135

LIST OF TABLES

CHAPTER 2		25
1.	<i>Table 1: Description of sample</i>	39
2.	<i>Table 2: Pearson correlation coefficients of CPT with validation measures</i>	41
3.	<i>Figure 1: ROC curve for CPT with SMMSE as criterion</i>	42
4.	<i>Table 3: Phi test: CPT score and SMMSE cut-off</i>	43
5.	<i>Table 4: Phi test: CPT score and AMPS-P cut-off</i>	43
CHAPTER 3:		55
1.	<i>Table 1: Test-retest reliability and validity of Incidents of Harm Questionnaire</i>	66
2.	<i>Table 2: Validity comparison between caregiver and medical chart on type of incident</i>	68
CHAPTER 4:		76
1.	<i>Table 1: Comparison of participants who did and did not experience harm</i>	89
2.	<i>Table 2: Sensitivity and specificity of functional measures for incident of harm</i>	91
3.	<i>Table 3 Cox regression analysis predicting time to harm: single variable analysis (n=45)</i>	92
4.	<i>Table 4: Cox regression model: contribution of independent variables in predicting time to harm with AMPS-P (n=45)</i>	94
5.	<i>Table 5: Cox regression model: contribution of independent variables in predicting time to harm with AMPS- M (n=45)</i>	95
6.	<i>Table 6: Cox regression model: contribution of independent variables in predicting time to harm with CPT (n=45)</i>	96
7.	<i>Table 7 Covariance matrix: Pearson's r correlation of all variables</i>	98

LIST OF FIGURES

CHAPTER 1:	1
1. <i>Figure 1: Consent process flow chart</i>	22
CHAPTER 2.....	25
2. <i>Figure 1: ROC curve for CPT with SMMSE as criterion</i>	42
CHAPTER 3:	55
3. <i>Figure 1: Study procedures</i>	63

LIST OF APPENDICES

CHAPTER 3:	55
<i>Appendix 1</i> :	75
CHAPTER 4:	76
<i>Appendix 1</i> :	113
CHAPTER 5.....	114
<i>Appendix 1</i>	138
<i>Appendix 2</i>	139
<i>Appendix 3</i>	142

CHAPTER 1: INTRODUCTION

Research Problem

For many older adults who are admitted to hospital, decisions about being able to live at home are difficult. Decisions are based on weighing the risks and benefits of returning home, and considering individual autonomy. Occupational therapists assist with this decision-making by assessing the patient and providing recommendations about safety at home. Although measures of daily function and cognition are often used to make inferences about home safety, little is known about their ability to predict safety accurately. The research described in this thesis arose out of the clinical need to use in-hospital assessment to predict safety after discharge and the lack of data to support the use of existing measures for this purpose.

Clinical Context and Background

For the purposes of this research, safety is defined as the avoidance of harm or injury to both person and property (Tierney, 2003). Although the home may be considered a place of refuge and comfort, concerns about safety for older adults are important in making decisions about the amount of paid or unpaid help required, or moving to supported living or long-term care. The concept of safety is important in discharge decision-making because patients and caregivers want and need to be fully informed about the risks to the individual. The basis for this is two-fold: one is

the ethical principle of planning discharge without causing harm; the other is to use resources effectively, such as allocating the correct amount of home care or preventing re-admissions. Typically, in the clinical setting of the hospital, the discharge planning process attempts to ensure a discharge plan that maximizes the individual's safety while still promoting autonomy. Clinical assessment by the health care team includes the identification of potential sources of risk, which are then reviewed with the patient and caregiver. The health care team may include physicians, nurses, social workers, dietitians, physical and occupational therapists.

The clinical context in which assessments are used is one that requires evidence-based recommendations. Assessment results are considered in combination with other sources of information, from family and health care professionals. This is important to note because of ethical considerations regarding interpretation of standardized test scores. The information supports autonomous decisions, and is not intended to remove decision-making capacity from a person who is deemed capable of making medical decisions.

Occupational therapy role with persons with dementia

Occupational therapists' core domain of interest is occupation, which is defined as activities and tasks of everyday life including self-care, productivity and leisure (Townsend, Polatajko et al., 2007). The Canadian Model of Occupational Performance (CMOP) (Canadian Association of Occupational Therapists, 1997) and its extension, the Canadian Model of Occupational Performance and Engagement

(CMOP-E) (Townsend, Polatajko et al., 2007) illustrate the main areas of assessment and intervention for occupational therapists, which include: occupations, person and environment. The scope of occupational therapy is described in this model to encompass aspects of the person and environment that are related to occupation. In the clinical context of the hospital, occupational therapists directly observe person factors that affect occupation including physical, affective and cognitive performance components. Information about the environment such as social, institutional, physical and cultural components is commonly gathered through interviews and discussion with the patient and caregivers. Factors that affect home safety are addressed by occupational therapists in terms of the ability to perform occupations at home. For the population of persons who are admitted to hospital and making decisions about whether or not they can safely continue to live at home, the most important occupations are activities of daily living (ADLs) (essential self-care tasks e.g. bathing, dressing, grooming) and instrumental activities of daily living (IADLs) (a secondary set of tasks required for independent living e.g. cooking, shopping)(Lawton, 1971). Safety concerns arise regarding the ability to manage daily tasks and occupational therapists commonly assess cognitive capacities as a “performance component” of daily occupations.

The assessment process also typically includes an interview to obtain self-report information about the home environment. Physical and cognitive capacity information obtained through self-report is typically corroborated with direct observation by the therapist. Although advantageous for ease of administration and

low cost, for persons with dementia, the disadvantage of self-report measures is the lack of certainty that the person comprehends and recalls the information (Coman & Richardson, 2006). Proxy-report is often used for persons with dementia, which commonly constitutes a caregiver interview. Proxy reporting is accurate for observable phenomena such as physical mobility and activities of daily living but less accurate in issues such as pain according to the findings of a systematic review (Oczkowski & O'Donnell, 2010). In an inpatient rehabilitation unit, the proxy caregiver interview is generally obtained as supporting information for direct observations of the patient using performance measures.

Occupational therapists' use of cognitive assessment measures:

The occupational therapy literature describes theoretical approaches that therapists use to assess cognition that are commonly divided into “bottom-up” and “top-down” approaches. Using the “bottom-up” assessment approach (Duchek & Abreu, 1997; Grieve, 2000; Vining Radomski 2002), a therapist focuses on cognitive capacities, such as memory or attention, often using pen and paper or other table top assessments, using performance to infer potential function in daily life. The advantage of this approach is that it reduces therapist bias, by depending less on clinical observations (Vining-Radomski, 2002) and provides quantifiable data that are easier to communicate (Grieve, 2000). The disadvantage is that it relies on the assumption that one cognitive construct can be assessed separately from another

(Vining-Radomski, 2002), and is not affected by factors such as unfamiliar assessment environment, education, language and culture.

The “top-down” approach (Duchek, & Abreu, 1997; Grieve, 2000; Vining Radomski, 2002) relies on the therapist's observation of performance on everyday tasks to ascertain cognitive capacities. The term “top-down” is used in the rehabilitation literature to describe a theoretical approach to intervention planning which also emphasizes examination of social and occupational roles to address relevant skills and impairments (Gordon & Quinn, 2003). The “top-down” approach has the advantages of allowing assessment in a more natural context, reflecting daily life more accurately (Grieve, 2000), and minimizing the effects of communication deficits (Vining-Radomski, 2002). These assessments also allow the therapist to observe the interaction of various cognitive skills as the person carries out daily tasks. A disadvantage is that it is difficult to standardize top down assessments: they can be easily influenced by the therapist's definition of “normal performance”, and ability to make clinical observations (Vining-Radomski, 2002). As well, the cognitive performance components may be more difficult to identify (Grieve, 2000).

Evidence regarding occupational therapists' use of cognitive assessment measures contributed to the development of this study. A national survey of occupational therapists described current practice and highlighted the need for evidence to support functional measures (Douglas, Liu, Warren & Hopper, 2007). Therapists (n= 247) reported using a wide array of measures (n= 75 measures). Cognitive assessment measures were used frequently (more than once a month by

65% of respondents) and across various work settings including inpatient settings (52%), clients' homes (25%), long-term care (14%), clinics/private practice (5%) and mental health centres (4%). Choice of measures showed regional variation, with measures popular in one region not being used in another. Theoretical approaches ("bottom-up" and "top-down") were used similarly across the country. Respondents who reported choosing "top-down" assessments, despite choosing non-standardized assessments, chose them for the purpose of predicting function. The data indicated that the development of standardized assessments that predict function was needed.

A critical review of standardized measures for dementia in older adults was undertaken (Douglas, Letts & Liu, 2008). The measures were examined for data on reliability and validity and rated according to a rating scale for evidence on health care measurement (Law, 2003). Measures that were designed to predict function lacked research demonstrating predictive validity. The review did not find data regarding the prediction of incidents of harm for any measures used in occupational therapy.

The critical review found a discrepancy between the needs of clinicians for measures to predict function and the lack of supporting evidence for predicting function. Data are needed to support common clinical practice. Clinicians require data on whether there is a link between standardized measures and outcomes in the community. Given the need for further evidence to support the predictive ability of measures used with older adults, the following section describes the types of evidence that are needed to support the use of a measure.

Types of evidence to support a measure

The process of development of a health measurement scale involves gathering evidence regarding reliability, validity, acceptability and feasibility (Streiner & Norman, 2003). Each of these attributes of a measure is described in relation to the clinical problem identified in this research.

Reliability is the degree to which confidence can be placed in the reproducibility of the test results (Streiner & Norman, 2003). As a result, reliability places an upper limit on validity (defined here as whether or not the test measures what it is intended to measure). A test is reliable if it provides consistent results and, thereby, is associated with little measurement error. For clinical measures, important types of reliability include test-retest, inter-rater, intra-rater, and internal consistency reliability.

Test-retest reliability is the consistency of test results with the same rater and subject on two separate occasions. It is relevant when one anticipates the construct being tested to remain stable over time. The degree of consistency is determined using Pearson's r correlation or intra-class correlation coefficient (ICC). Pearson's r is a more liberal measure because it describes the extent to which the scores can be fitted to a straight line without accounting for the intercept or slope of the line (Streiner & Norman, 2003). However, often the largest component of variation is attributable to random factors and, in such situations, Pearson's r will return values similar to the ICC. For measures with dichotomous responses (yes/no), test-retest reliability can also be calculated using kappa though kappa and

ICC are mathematically identical (Fleiss and Cohen, 1978). Sources of error particularly relevant to test-retest reliability include random variation (noise) and non-random variation which can include changes in testing circumstances (e.g. day of the week) and fluctuations in the state of the subject on other factors that may affect test performance such as fatigue or stress.

Inter-rater reliability is the consistency of test results obtained from the same subject by two separate raters. Intra-rater reliability describes consistency of a single rater across time when the stimulus is held constant. Steps to increase inter- and intra-rater reliability of a test include development of clear administration and scoring instructions, and rater training. Internal consistency reliability is the degree to which each item correlates consistently with other items on the test. It gives an indication of the consistency with which the items in the measure reflect a uniform construct.

Validity is the degree of confidence that can be placed in the interpretation of the scores, and as such is a fundamental consideration in the use of tests in health care (Streiner & Norman, 2003). Concepts of validity have been evolving. Classical test theory describes validity by listing three major types: construct (degree of association with theoretical concepts), content (degree of representativeness with the target concept) and criterion (correlation with another measure of the construct, ideally a “gold standard”)(Streiner & Norman, 2003). Kane describes the difficulties that are encountered when applying classical test theory as twofold: the criterion against which a test is validated may itself be flawed or the construct against which

the construct is tested may not be relevant. Messick and Kane argued that ultimately, the concept of validity needed to evolve because the three categories set up the misconception of validity as a checklist. Validity was viewed as a static property of a test that once studied did not require further consideration of different settings, populations and social consequences to the purpose to which the test was used. However, the concepts of construct, concurrent and criterion validity remain valuable to demonstrate validity in the early stages of test development (Kane, 2006).

Subsequently, validity has been defined as a single concept, to which multiple types of evidence contribute. Messick (1995) argued that the most important evidence is that which validates the way the test is interpreted; that is, the consequences and social impact of the test score. Studies must be designed that evaluate the outcomes that are used in interpretation. The concepts of content, construct, and criterion validity are encompassed within this conception of validity. This conception of validity has been incorporated into the Standards for Educational and Psychological Testing which list the types of evidence that contribute to overall validity as evidence based on: test content, response process, internal structure, relations to other variables, and consequences of testing (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999). These standards recommend that validation must be shown for each intended interpretation of a score.

Most recently, validation is described as requiring scientific argument using multiple lines of evidence (Kane, 2006). “Argument-based” validity, as Kane (2006) describes it, is seen as an ongoing scientific process. The process of validation must employ two kinds of arguments. The first is the interpretive argument for which a statement is made about how a test is to be interpreted, including assumptions and inferences. The second is the argument for validity. It demonstrates that the interpretation argument is coherent (i.e., the inferences are reasonable and assumptions are plausible). A variety of studies are required to support each specific use of the assessment. Thus, a particular research design for validation is not prescribed because the design needs to match the way the test is interpreted. The validation study needs to form the strongest argument for the test interpretation.

Kane describes important components for a coherent validation argument. It includes a rationale for why the measure is expected to relate to the outcome, including the theoretical construction of the tool, and further testing of the measure. Further testing itself can include analyzing scoring, generalization, extrapolation, or decision rules (Kane, 2006). Analyzing scoring means whether the scoring rule is applied accurately and consistently. Generalization means whether the observations are representative and numerous enough to control sampling error and extrapolation is the ability of the score to represent the overall attribute or skill. Finally, decision rules are whether the score on the measure accurately corresponds

to a decision about the characteristics being measured (e.g. “impaired” or “unimpaired”).

The final measurement characteristics, acceptability and feasibility, relate to perceptions on the part of the test administrator and test recipient (Streiner & Norman, 2003). These characteristics may affect the uptake of a test in practice irrespective of whether a test is supported by evidence of reliability and validity. Acceptability refers to the perception of the burden of testing. Tests that require what is perceived as excessive effort or are invasive by causing hardship, embarrassment, pain or distress may be less likely to fully engage the person who is tested and may generate inaccurate measures of the construct being measured. Feasibility refers to the perception of the fit with clinical practice. A test that is regarded to take excessive time and materials, to be of high cost, or to require set-up that does not work in the setting, is less feasible and less likely to be used in clinical practice regardless of reliability or validity. If measures have similar reliability and validity evidence, generally one would expect the most feasible measure to be more popular and useful clinically.

Study design and measure selection

The clinical problem to be addressed is the validation of in-hospital assessments for predicting harm outcomes in order to inform the discharge planning process. Given the above background regarding reliability and validity, it was important to choose measures with established reliability, and design a validity

study to determine if the measures are associated with clinically important outcomes. A prospective design is the optimal research design to provide a line of evidence regarding harm after hospital discharge. This is because our focus is on the relationship between the scores obtained and future events irrespective of the factors that contributed to the outcome to take into account all factors that may influence the outcome. The STROBE statement for the design of prospective cohort studies (Vandenbroucke, et al., 2007) was consulted in the development of the validation study. The STROBE checklist includes defining the population and approach of all eligible persons, ensuring outcome measures have supporting psychometric data, minimizing burden for long-term follow-up, and monitoring for loss to follow-up.

The concept of validation for the exact interpretation of a test result was important for the development of this study. The outcome chosen (harm experienced after hospital discharge) was chosen as the most clinically important outcome to match with the way the test is being interpreted clinically. Thus, the concept of validation by Kane supports the design of this study. During the formative stage of this study, the concepts of validity as described by Kane (2006) had not disseminated into the rehabilitation field. Because the manuscripts were being written for a rehabilitation audience, classical test theory terms such as criterion validity are used. The use of these terms does not exclude the work of Messick (1995) or Kane (2006). They are used to describe how the evidence can be used to inform interpretation of the measure for the clinical reader.

The rationale for the measures chosen for the validation study included both psychometric evidence and feasibility. First, the psychometric data were examined to support various cognitive measures used in occupational therapy for persons age 65 years and over (Douglas, Letts & Liu, 2008). Amongst the measures reviewed, the Assessment of Motor and Process Skills (AMPS; Fisher, 2006) had several advantages by being a measure uniquely designed to assess daily occupations by an occupational therapist, using a “top-down” theoretical approach, and had the highest rigour of psychometric data to support its use.

The AMPS was designed to observe the performance of daily living tasks of the client's choice, for a wide range of ages and diagnostic groups. The therapist selects six or eight tasks from a possible list of at least 80 tasks, and speaks with the client to select the two most relevant and meaningful tasks that they know how to perform. The client is then observed performing the two tasks (approximately one hour) and performance is scored on 15 motor (e.g. lifts, walks) and 20 process (e.g. initiates, sequences) skills to obtain two overall scores: motor and process score. The rater scores the observable performance and does not assume which underlying capacities affect performance. Scores are converted to a relative ranking using Rasch analysis, which takes into account task difficulty and rater severity. A cut-off score delineating a recommendation of “independent” or “dependent” is provided. Reliability data are available for test-retest, inter-rater and internal consistency and were rated as “excellent” based on availability of data from two or more studies (Douglas, Letts & Liu, 2008). The test manual cites validity data that have

demonstrated discrimination between persons with and without disability in two studies, including persons with dementia (Fisher, 2006). It has been correlated with at least five other ADL or cognitive measures, and demonstrated cross-cultural validity in two studies (Fisher, 2006). Evidence shows association with independence in the home (McNulty & Fisher, 1999), and more recently, driving ability (Dickerson, Reistetter & Trujillo 2010). Administration requires occupational therapy qualifications, AMPS training and certification. Part of the preparation for this research study involved completion of the AMPS training and certification by the principal investigator.

The second measure was chosen based on the argument of feasibility. The Cognitive Performance Test (CPT) (Burns, 2006) had data supporting its psychometric properties, the advantages of being less costly to purchase (approximately \$200 USD versus \$750 USD for AMPS), and because it was part of usual care at the proposed research site.

The CPT (Burns, 2006) was designed to observe the performance of persons with dementia on specific ADL and IADL tasks that require information processing skills. The individual performs seven activities (Medbox, Dress, Shop, Toast, Phone, Wash, and Travel)(approximately 45 minutes). Scoring is based on Allen Cognitive Level Theory for the level of cueing and demonstration required (score 1 to 6). Scores on the seven tasks are averaged to determine the total score. To ensure that the total score is not skewed, it is important that at least 4 subtasks are completed and that subtasks are not individually assessed so that a global functioning score is

analyzed (Burns, 2006). Descriptors based on theoretical understanding of the stages of dementia are provided for each level to aid in score interpretation. They provide interpretation for the clinician that describes the amount of cueing that will likely be required by the person during daily living. This interpretation is at the functional level as described by Nagi (1965), which is the ability to perform specific tasks. High test-retest, inter-rater reliability and internal consistency reliability was found in two studies with persons with Alzheimer’s disease (Burns, Mortimer, & Merchak, 1994; Bar-Yosef, Weinblatt, & Katz 1999). This evidence was rated to be “adequate”, because it is based on data from two studies (Douglas, Letts & Liu, 2008). Validation data include one study discriminating between normal older adults and those with Alzheimer’s disease (Bar-Yosef et al., 1999) and two studies correlating CPT scores with cognitive and ADL measures (Burns et al., 1994; Thralow & Reuter, 1993). Low scores were predictive of institutionalization after 624 days (Burns, et al., 1994). The test was designed for administration by an occupational therapist, and does not require specialized training. Of the two measures, CPT and AMPS, the CPT would benefit from further published data describing its validity because there are fewer publications to support its psychometric properties and its easy accessibility to therapists means that evidence is needed to inform its use.

Thus, the measures were chosen because they showed promise for in-hospital evaluation of older adults by occupational therapists. Both were appropriate for pre-discharge evaluation and had data to demonstrate their

potential for predicting function after discharge. As measures designed to indicate levels of independence in daily function, it was hypothesized that a link to injury or harm might be made. Thus, the validation study was designed to examine the validity of the measures related to a new outcome that is clinically important because it that takes into account the performance in the real-world, rather than a proxy measure of independence. Examination of the validity of AMPS and CPT for predicting harm could give important additional information to aid in the interpretation of scores. Ethical use of the measures requires that evidence be provided to either support or refute this interpretation.

Study Purpose and Objectives:

The overall purpose of this study was to examine whether two functional measures were predictive of harm in the community. The study examined the use of the measures in a population of older adults who were admitted to the hospital to determine if scores obtained as an inpatient were predictive of outcomes after discharge. The objectives of the three manuscripts produced from the study were to a) describe the validity of the Cognitive Performance Test (CPT) using baseline data b) examine the reliability and validity of the “incident of harm” outcome used to measure harm, and c) to examine the correlation of the AMPS and CPT with the “incident of harm” outcome as well as the contribution of other independent variables to predict the outcome. A brief summary of these manuscripts follows.

The first manuscript (Chapter 2) is entitled “Use of the Cognitive Performance Test (CPT) to identify deficits in hospitalized older adults”. Hereafter this study will be called “the CPT study”. This paper builds on previous studies of the CPT in establishing its ability to measure cognition independently of other factors. Previous studies have shown concurrent validity with other measures of cognition but it is not known if the CPT measures cognition independently of age, sex and education level. Also, it is not known if the CPT measures the construct of cognition separately from other constructs such as daily living skills, chronic medical conditions and motor skills. Moreover, it has not been determined whether, compared to other measures of cognition, the CPT identifies persons who require assistance to live in the community. The objectives of this manuscript were to: 1) determine whether age, sex and years of education correlate with CPT scores, 2) determine discriminant validity of the CPT compared to measures of ADL (Functional Independence Measure (FIM)), chronic medical illness burden (CIRS-G), and motor skills (motor scale of the AMPS), 3) determine concurrent validity with measures of cognitive screening (Standardized Mini Mental Status Exam (SMMSE)), and process skills (process scale of AMPS), and 4) determine the agreement of CPT with SMMSE and AMPS regarding decision rules defining the threshold for impairment. The data from this study comprise the baseline data from the prospective study described in the thesis. The relevance of the study is to describe the degree of confidence that clinicians can place on determination of cognitive impairment with the CPT, and interpretation of functional status (Nagi, 1965).

Determination that the CPT score does not correlate with age, education or comorbidities would support the construct validity of the measure as a measure of cognition separate from daily living skills, co-morbidities and motor skills.

Furthermore, if it designates impairment similarly to other measures, this provides evidence for the ability of the CPT to accurately identify persons with impairment.

The second manuscript (Chapter 3) is entitled “Measurement of harm outcomes in older adults after hospital discharge: reliability and validity”, and will hereafter be called “the harm outcome study”. This manuscript describes the clinical importance of the need to measure harm outcomes to validate clinical measures.

The objectives of the manuscript were to determine the test-retest reliability and validity of measuring the outcome “incidents of harm” by caregiver telephone interviews. The measurement method built on the method described previously for measurement of harm as an outcome (Tierney et al., 2004). The measurement of harm outcome is relevant to researchers for validation of clinical assessment tools and telephone interviews are a highly feasible way of collecting such information.

This outcome/response process is used in the third manuscript to examine the validity of two functional measures of cognition in occupational therapy.

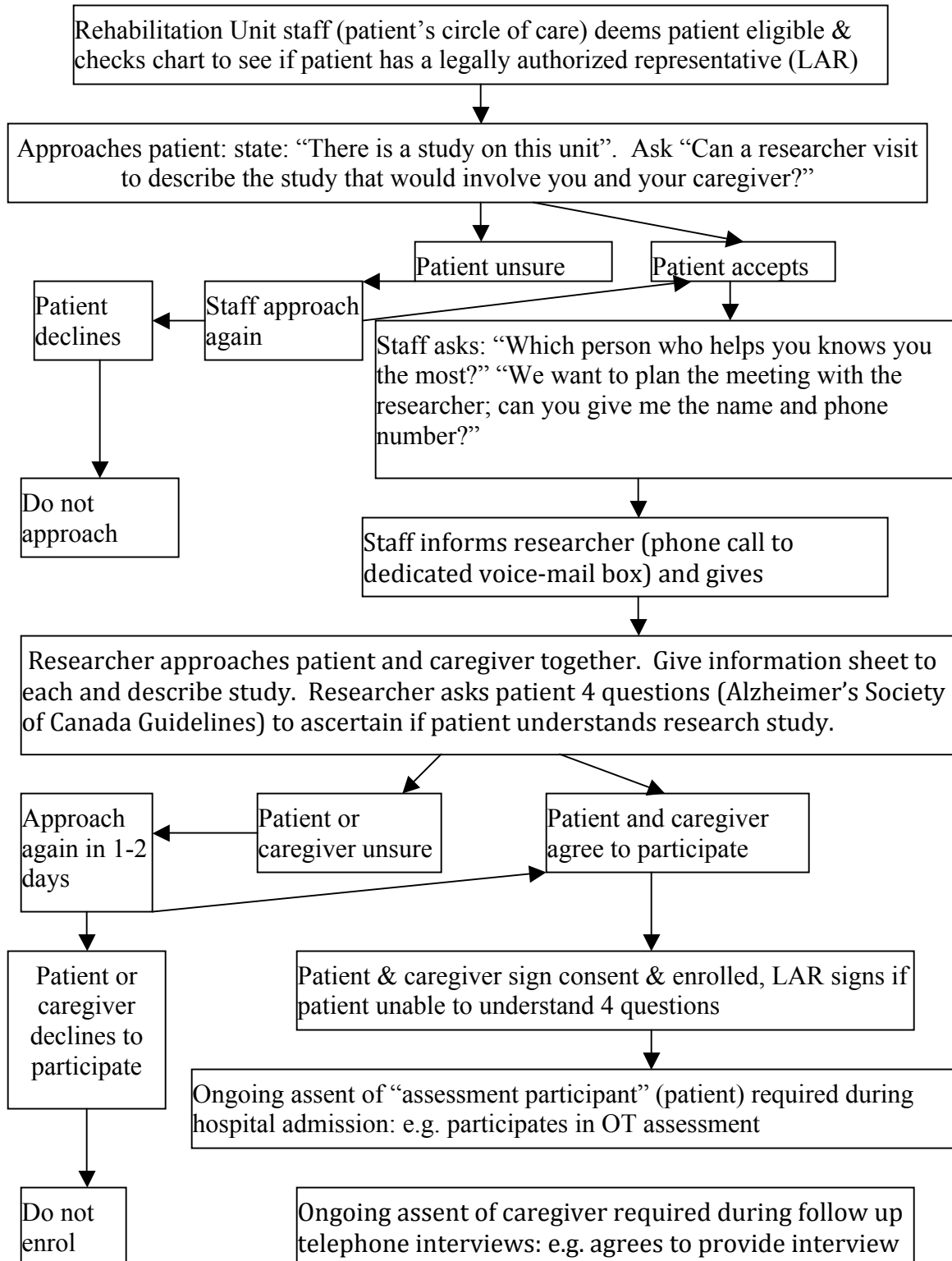
The final manuscript (Chapter 4) is entitled “Validity of pre-discharge functional measures for predicting harm in older adults”. The purpose of this manuscript was to examine whether two pre-discharge functional measures were valid for predicting incidents of harm after discharge. It will hereafter be called “the predictive validation study”. Effective discharge planning for older adults includes

the consideration of safety and potential for harm. Therefore, this study sought to validate assessment measures for determining risk for harm. The specific objectives of the study were to determine 1) the ability of CPT score to predict harm outcome, as measured by the incidents of harm reported over a 6-month community follow-up, 2) the ability of AMPS- Motor score and AMPS- Process scores to predict the harm outcome, as measured by incidents of harm reported over a 6 month community follow-up and 3) the contribution of independent variables (age, sex, education, comorbidities, activities of daily living independence score, and caregiving hours) to predict the harm outcome. The importance of this study is that it provides evidence to support the interpretation of the scores on these measures for discharge planning, in comparison with other variables, and offers data to identify which factors might increase the risk for incidents of harm after discharge.

Several ethical issues arise when considering this type of research design. Persons with dementia are considered a vulnerable population. It is, therefore, particularly important to take measures to reduce the possibility of coercion while also encouraging participation in the research study. The most feasible solution was to require consent be obtained from both the hospital patient and the caregiver proxy, and for the consent to be obtained during the same meeting. In this way, both could hear the same information and make a decision together. Furthermore, during the consent process the patient was assessed for his or her ability to understand the basic purpose, risks and benefits of the study so that the substitute decision maker could consent on behalf of the patient if needed. The steps required during

recruitment and obtaining consent were sufficiently complex that the best way to clearly describe them was to create a flow chart (Figure 1). This diagram was provided for the ethics review process and also enabled the hospital staff and research assistants to maintain ethical integrity of the study recruitment process. Also, the ethics board required documentation of the clinical context in which the assessments were interpreted to ensure that the standardized scores for the usual care measures were used to contribute to autonomous decision making about long term care and discharge planning.

Figure 1: Consent process flow chart



Summary:

In this introductory chapter the reader is provided with an overview of the clinical problem in occupational therapy related to the use of measures for discharge planning. It provides context by describing the clinical population of persons with dementia and measurement theory. The rationale for research design and choice of measures for the study are described. Each of the three manuscripts was summarized in terms of background, purpose and relevance. There may be some overlap between the content of each manuscript because each is written for independent submission to separate publications. There was need to include all relevant background although it may have been included in a previous paper. Following the three manuscripts is a discussion chapter including the contribution of each manuscript and the overall contribution of the research both academically and clinically.

References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for Educational and Psychological Testing*. Washington, DC: American Psychological Association.
- Bar-Yosef, C., Weinblatt, N., & Katz, N. (1999). Reliability and validity of the cognitive performance test (CPT) in an elderly population in Israel. *Physical and Occupational Therapy in Geriatrics*, 17(1), 65-79.
- Burns, T. (2006). *Cognitive Performance Test Manual*. New Jersey: Maddak.
- Burns, T., Mortimer, J. A., & Merchak, P. (1994). Cognitive performance test: A new approach to functional assessment in Alzheimer's Disease. *Journal of Geriatric Psychiatry and Neurology*, 7, 46-54.
- Canadian Association of Occupational Therapists. (1997). *Enabling occupation: An occupational therapy perspective*. Ottawa, ON: CAOT Publications ACE.
- Coman, L. & Richardson, J. (2006). Relationship between self-report and performance measures of function: a systematic review. *Canadian Journal on Aging* 25, 253-270.
- Dickerson, A. Reistetter, T., Trujillo, L. (2010). Using an IADL assessment to identify older adults who need a behind the wheel driving evaluation. *Journal of Applied Gerontology* 29, 494-506.
- Douglas, A., Letts, L., & Liu, L. (2008). Review of cognitive assessments for older adults. *Physical and Occupational Therapy in Geriatrics* 26(4), 13-43.
- Fisher, A. G. (2006). *Assessment of Motor and Process Skills. Vol. 1: Development, standardization & administration manual* (6th ed.) Fort Collins, CO: Three Star Press.
- Gordon, J. & Quinn, L. (2003). *Functional Outcomes: Documentation for rehabilitation*. St Louis, Missouri: Elsevier Science.
- Kane, M. T. 2006. Validation. In R. Brennan (Ed.) *Educational Measurement, 4th edition*. Wesport, CT: American Council on Education and Praeger, 17 - 64.
- Law, M. (Ed.). (2003). *Evidence-based rehabilitation: A guide to practice*. Thorofare NJ: Slack Incorporated.

- Lawton, M.P. (1971). The functional assessment of elderly people. *Journal of the American Geriatrics Society*, 19, 465- 481.
- Messick, S. (1995). Validity of psychological assessment. *American Psychologist* 50, 741-749.
- Nagi, S.Z. (1965). Some conceptual issues in disability and rehabilitation. In M.B. Sussman (Ed.), *Sociology and rehabilitation* (pp. 100–113). Washington, DC: American Sociological Association.
- Oczkowski, C. & O'Donnell, M. (2010). Reliability of proxy respondents for patients with stroke: a systematic review. *Journal of Stroke and Cerebrovascular Disease* 19, 410-416.
- Streiner, D.L., & Norman, G.R. (2003). *Health Measurement Scales: A practical guide to their development and use* (3rd ed.). New York: Oxford University Press.
- Thralow, J., & Rueter, M. (1993). Activities of daily living and cognitive levels of function in dementia. *American Journal of Alzheimer's Care and Related Disorders and Research*, 8, 14-19.
- Tierney, M. C., Charles, J., Naglie, G., Jaglal, S., Kiss, A., & Fisher, R. H. (2004). Risk factors for harm in cognitively impaired seniors who live alone: A prospective study. *Journal of the American Geriatrics Society*, 52, 1435-1441.
- Townsend, E.A. & Polatajko, H. J. (2007). *Enabling Occupation II: Advancing an Occupational Therapy Vision for Health, Well-being & Justice through Occupation*. Ottawa, ON: CAOT Publications ACE.
- Vandenbroucke, J.P., von Elm, E., Altman, D.G., Gøtzsche, P.C., Mulrow, C.D., Pocock, S.J., Poole, C., Schlesselman, J.J., & Egger, M. (2007). STROBE Initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and elaboration. *PloS Medicine*, 4 (10) e297, 1628-1654
- Vining Radomski, M. (2002). Assessing abilities and capacities: cognition. In C.A. Trombly & M. Vining-Radomski (Eds.), *Occupational Therapy for Physical Dysfunction 5th edition*. (pp.199-212). Baltimore: Lippincott Williams & Wilkins.

CHAPTER 2

**Use of the Cognitive Performance Test (CPT) for identifying deficits in
hospitalized older adults**

Authors:

Alison Douglas

Lori Letts

Kevin Eva

Julie Richardson

Target journal: American Journal of Occupational Therapy

**Use of the Cognitive Performance Test (CPT) for identifying deficits in
hospitalized older adults**

Abstract

Objectives: The Cognitive Performance Test (CPT) is a functional assessment for persons with dementia. The study purpose was to contribute validation evidence for the CPT. Namely whether: (1) age, sex, education motor skills and co-morbidities influenced the score, and (2) it correlated with other measures on designation of impairment.

Method: The CPT was tested against other measures of cognition (Standardized Mini Mental Status Exam (SMMSE) and Assessment of Motor and Process Skills- Process scale (AMPS-process)). Participants were persons 65 years and older admitted to a geriatric rehabilitation unit (n=47).

Results: The CPT correlated moderately with measures of cognition (SMMSE $r=0.47$, AMPS-P $r=0.53$, $p<0.01$), and ADL burden of care (FIM $r= 0.32$, $p<0.05$). Scores were not affected by age, sex, years of education, motor skills or co-morbidities. The CPT differentiated impaired and unimpaired differently from other measures.

Conclusion: While CPT appears related to other measures of cognition, test interpretation requires noting the variability between CPT scores and those measures.

Introduction

Cognitive impairment is one of the strongest predictors of institutionalization, increasing the risk by two and a half times (Gaugler, Duval, Anderson & Kane, 2007).

There is likely to be an increase in institutionalization with the aging population and an associated individual and societal burden. Evidence based care is a priority with this patient population for caregivers, families, individuals and policy makers.

Dementia is a neurological deficit characterized by memory loss and other cognitive functions (American Psychiatric Association, 2000, *Diagnostic and statistical manual of mental disorders* (text revision)). Both of The following criteria need to be confirmed for a classification/ diagnosis of dementia: 1) decline in memory and at least one of the following: language, object recognition, motor skills, abstract thinking, judgments and planning and executing complex tasks. 2) decline must be severe enough to interfere with daily life (*DSM-IV-TR*, 2000). There are various types of dementia including Alzheimer's disease, vascular dementia, and mixed dementias. The prevalence is 6.9 % in North America and the worldwide incidence is projected to double in ten years (Alzheimer's Disease International, 2009). The role of occupational therapy is to enable the performance of everyday activities and participation (Crepeau, Cohn & Schell, 2003; Townsend & Polatjko, 2007). For persons with dementia, there is impairment in the ability to perform activities of daily living (ADL) (essential self-care tasks e.g. bathing, dressing, grooming) and instrumental activities of daily living (IADLs) (a secondary set of tasks required for independent living e.g. cooking, shopping)(Lawton, 1971). Occupational therapists who see people with dementia in hospitals are concerned with their ability to live independently and safely after discharge and to recommend the amount of assistance required (Bonder & Wagner, 2001; Vining Radomski,

2002). Occupational therapists need to use an evidence-based evaluation the ability of persons with dementia to complete functional tasks.

The Cognitive Performance Test (CPT) was developed for the assessment of persons with dementia (Burns, 2006). It is designed to structure observations of the performance of persons on specific ADL and IADL tasks that require information processing skills. The test is designed for administration by an occupational therapist (OT), and does not require specialized training. Test interpretation is intended to assist with treatment planning and to predict a person's need for assistance. The attributes of feasibility of cost, level of training, and administration time means it has potential for uptake amongst clinicians.

Occupational therapy literature describes two primary approaches to assessing cognition: a "bottom-up" approach and "top-down" approach (Duchek, & Abreu, 1997; Grieve, 2000; Vining Radomski, 2002). A "bottom-up" approach examines cognitive components to infer a person's function. The CPT uses a "top-down" approach, relying on the observation of performance on everyday tasks to ascertain cognitive abilities. The CPT is not designed as a measure of skill on specified tasks (e.g. ability to manage medication) but as a measure of working memory in everyday function (Burns, 2006). The author of the CPT defines working memory as an important cognitive component for multiple cognitive processes including focused attention, selective attention, response inhibition, task switching, abstraction, formulating goals and plans, adjusting behaviour, problem solving, decision making, reasoning and judgment. Therefore working memory is argued to

be needed for carrying out complex tasks such as instrumental activities of daily living (Burns, 2006).

The CPT score is interpreted according to cognitive-functional profiles. These cognitive functional profiles are derived from a theoretical understanding of neurocognition as described by Burns and Levy (2005) and are included in the appendices of the test manual. The profiles describe differing levels of working memory function and the authors state that deficits in working memory constrain function. Burns and Levy (2005) describe six “cognitive levels” based on Allen’s Cognitive Disability Theory (Allen, 1990 as cited in Burns, 2006). Starting at level six (normal), each successively lower “cognitive level” (5,4,3 etc.) is described according to function in daily living. Each level has an accompanying “cognitive-functional profile” which is described in the test appendix. This profile is intended to assist the therapist with intervention planning and education about caregiving strategies.

The evidence to support the reliability of the CPT is based on four studies. Test-retest reliability, over a four-week interval was reported as Pearson’s $r=0.89$ ($n=36$) (Burns, 1992). Inter-rater reliability between two raters was $r=0.91$ ($n=18$) and $kappa=0.98$ between two raters of two clients on video (Bar-Yosef et al., 1999). Internal consistency reliability was high ($r=0.84$ Burns, 1992; 0.95 Bar-Yosef et al., 1999). A review comparing measures for older adults rated this reliability evidence as ‘adequate’ (Douglas, Letts & Liu, 2008) based on there being only one to two well-designed studies for each type of reliability (Law, Baum & Dunn, 2005).

The CPT has been validated using concurrent validity with other measures of cognition and activities of daily living (ADL). Validity data for the CPT is based on three published studies (Bar-Yosef et al., 1999; Burns, 1992; Thralow & Reuter, 1993), resulting in the evidence for validity being rated as 'adequate' (Law, Baum & Dunn, 2005). Concurrent validity was demonstrated with the MMSE (Pearson's $r = .67$, $n = 36$ (Burns, 1992); $r = .76-.88$, $n = 60$ (Bar-Yosef et al., 1999)) and ADL function using the Routine Task Inventory ($r = 0.91-0.96$ Bar-Yosef et al, 1999), or Self Care Performance Test ($r = 0.78$, Thralow & Reuter, 1993). It has been correlated with caregiver reported ADL function using the Routine Task Inventory-caregiver ($r = 0.50-0.68$ ($n = 60$) (Bar-Yosef et al., 1999). The CPT manual cites unpublished data to support a correlation with neuropsychological tests of planning, sequencing and attention (Bares, 1998)($n = 100$). Measures of episodic memory, language and co-morbidity were not significantly correlated with CPT scores (Bares, 1998). Predictive validity was described in a study that demonstrated that a cognitive level of 4.2 or lower was a significant predictor of institutionalization over 4 years of follow up (Burns, 1994).

Given this evidence, the CPT shows promise and would benefit from further study to address gaps in validity data. The CPT is intended to be a measure of cognition; therefore, it should be expected to correlate with measures of cognition but not with demographic characteristics or other constructs such as motor skill and ADL. It is not known if CPT measures cognition independently of demographic characteristics of age, sex and education level. If any of these affect the score, a

person who is designated impaired (or unimpaired) by the test, may actually be unimpaired (or impaired) compared to others their age, sex or education level, thereby reducing the validity of the test interpretation or requiring more normative data based on these sources of variation.

Also, it is not known if CPT measures the construct of cognition separately from other constructs such as daily living skills, chronic medical conditions and motor skills. The CPT author acknowledges that, although the test was designed to minimize the influence of motor skill, motor skill may influence scores and data are needed to examine this relationship (Burns, 2006). Determination that the CPT score does not correlate with the measures of daily living skills, chronic medical conditions and motor skills would support its construct validity. As a measure of cognition, CPT scores are expected to vary independently of daily living skills, comorbidities and motor skills.

Moreover, it has not been determined whether, compared to other measures of cognition, the CPT identifies persons who require assistance to live in the community. The test is designed to help identify persons who need assistance. Therefore it is important to understand the accuracy with which it identifies persons who are or are not impaired compared to other measures of cognition (designation of impairment). The precision of designation of impairment is a measurement property of the tool related to how well it differentiates persons who are deemed impaired versus unimpaired according to the tool's definition of impaired (Streiner & Norman, 2003). This could also be called the decision rule (Streiner & Norman,

2003). A previous study examined the designation of impairment for a measure similar to the CPT, called the Large Allen Cognitive Levels (Marom, Jarus & Josman, 2006). The Large Allen Cognitive Levels test is described by the CPT author as the basis for the development of the CPT and its scoring system (Burns, Mortimer & Merchak, 1994). The previous study found no relation between designation of impairment on the LACL and cognitive measures. Therefore, it is also important to determine whether the CPT designates impairment similarly to other measures of cognition. This evidence is needed clinically to determine validity of the CPT to correctly identify persons with impairment and in need of supports in the community setting.

This study aimed to examine the validity of the CPT by describing the degree of confidence that clinicians can place on CPT scores to identify cognitive impairment in older adults, and the interpretation of functional status in a hospital setting. The process of validation involves testing hypotheses to determine the balance of evidence regarding the extent to which the instrument appears to measure what is intended, thereby influencing the degree of confidence we can place on interpretation of the score in clinical practice (Streiner & Norman, 2003).

Validity data, together with characteristics of feasibility and acceptability are important to consider when choosing the most accurate and efficient assessments for persons with dementia. These assessment scores contribute to decisions around independent living, and therefore data are required to determine the influence of

age, education, co-morbidities, activities of daily living (ADL) status and motor skills on CPT scores.

The objectives of this study were to:

- 1) determine whether age, sex and years of education correlate with CPT scores
- 2) determine correlation of CPT scores with measures of
 - a) ADL burden of care (Functional Independence Measure (FIM))
 - b) Chronic medical illness burden (CIRS-G)
 - c) motor skills (motor scale of the Assessment of Motor and Process Skills (AMPS))
- 3) determine concurrent validity of CPT scores with measures of
 - a) cognitive screening (Standardized Mini Mental Status Exam (SMMSE))
 - b) process skills (process scale of AMPS).
- 4) determine the agreement of CPT with SMMSE and AMPS regarding decision rules defining the threshold for impairment.

It was hypothesized that the functional measure of cognition (CPT), would not be associated with age, sex, or years of education, nor would it strongly correlate with measures of ADL, chronic illness burden or motor skills. It was hypothesized that the CPT would show significant correlations with cognitive screening (MMSE) and process skills (AMPS- Process) and would define the threshold of impairment similarly to the SMMSE and AMPS-Process scale.

Methods

The study used a cross sectional design. The data for this study represented baseline measurements of a prospective study that examined the predictive validity of functional/cognitive measures to predict safety six months after discharge.

Participants

Participants were recruited from consecutive admissions to a geriatric rehabilitation unit at a large urban teaching hospital. Persons admitted to the geriatric rehabilitation unit were included in the study if they were 1) 65 years or older, 2) referred to occupational therapy, 3) English speaking, 4) undergoing functional or cognitive assessment with the OT because of a suspicion of cognitive impairment (diagnosis of dementia not required). Persons were excluded if they were 1) exhibiting symptoms of delirium, or 2) had a primary diagnosis of mental illness excluding dementia noted on the chart. Recruitment and consent processes followed ethics guidelines and were approved by the local health research ethics board.

Assessment Instruments

The Cognitive Performance Test (CPT) was part of usual care on the geriatric rehabilitation unit for this study. The test requires direct observation of an individual performing up to seven activities (Medbox, Dress, Shop, Toast, Phone, Wash, and Travel). Scoring is described for each subtest individually in the test manual. Overall, higher scores (5 or 6) are given when little cueing or demonstration is required and low scores (1) when cues, which are specifically described for each subtest, must be given for the patient to complete the task. Total administration time is approximately ½ hour and

test administration is made efficient by allowing elements of a subtask to be omitted if a person shows low functioning during the assessment. Directions for standardized cues are provided. Scores on each task are added and divided by the number of subtasks given to calculate an average task performance (total score). The total score is also converted into a “Cognitive Level” (lowest level represents the most impairment) that is based on theoretical understanding of function in stages of dementia (Burns, 2006).

The Standardized Mini Mental Status Exam (SMMSE) is a screening test for dementia with the same test items as the Mini Mental Status Examination (MMSE) (Folstein, 1975). The SMMSE differs from the MMSE because it has “clear, unequivocal guidelines for administration” (Molloy & Clarnette, 1999, p. 12) which increase the intra rater reliability compared to the MMSE (Molloy, Alemayehu, & Roberts 1991). The SMMSE has been reported to have greater accuracy than both generic and specific capacity assessment tools to identify incapacity to complete an advanced directive according to a gold standard (Competency Clinic) (Molloy, Silberfeld, Darzins, Guyatt, Singer, Rush, et al., 1996). The SMMSE is a widely used instrument and was part of usual care at the study site.

The purpose of the AMPS (Fisher, 2006) is to observe the performance of daily living tasks of the client's choice (approximately 15 minute interview), for a wide range of ages and diagnostic groups. The client is observed performing two tasks (approximately 1 hour) and performance is scored on 15 motor skills (e.g. lifts, walks) and 20 process skills (e.g. initiates, sequences). The AMPS scores incorporate four facets of assessment: item complexity, rater severity, task difficulty, and participant ability.

Ratings are entered into a computer program that uses Rasch analysis of the four facets to determine two independent overall scores: AMPS-motor and AMPS-process scores.

Each scale has a cut-off point above which indicates an independence level and below which indicates the need for assistance. Administration requires OT qualifications, a one-week training course and certification with ten patients to determine Rasch values for scoring. The data to support the psychometric properties of the AMPS are derived from numerous studies, which report excellent test-retest reliability, inter- and intra-rater reliability and validation with measures of ADL, IADL and cognitive tests (Douglas, Liu & Letts, 2008).

The Functional Independence Measure (FIM) involves observation by the clinician of the patient's level of independence in activities of daily living (Uniform Data System for Medical Rehabilitation, 2010) and was a mandated measure at the study site. The purpose of the measure is to describe the degree of disability and burden of care (Uniform Data System for Medical Rehabilitation, 2010). The patient's usual level of independence is observed during the daily routine. Scores are assigned from one (total assist) to seven (complete independence) on 19 daily living items (e.g. toileting, dressing) and are accumulated to determine the overall score, which indicates the degree of help required. For the total FIM score, five studies report high inter-rater reliability (ICC= 0.8-.99), and high internal consistency (alpha= .8-.97) (Glenny & Stolee, 2009). Validity has been described in 41 studies with older adults and have repeatedly shown correlation with assessments of daily living skills (Glenny & Stolee, 2009; Letts & Bosch, 2005;

McDowell & Newell, 1996) including in persons over age 80 (Pollak, Rheault & Stoeker, 1996).

The Cumulative Illness Rating Scale– for Geriatrics (CIRS-G) is a measure of chronic medical illness burden for older adults (Miller, Paradis & Reynolds, 1991). Diagnoses in 14 categories (e.g. vascular, renal) are assigned a severity score from 1(not severe) to 4 (severe). The index score is the total score divided by the number of co-morbid conditions and is an indication of the severity of concurrent conditions. Data demonstrate high levels of reliability (Miller, Paradis & Reynolds, 1991).

Procedures

The CPT was administered by the attending occupational therapist (OT) as part of routine care. The Standardized Mini Mental Status Exam (SMMSE) (Molloy & Clarnette, 1999), and the Functional Independence Measure (Uniform Data System for Medical Rehabilitation, 1994) were administered by unit staff as part of routine care. The participants' age, and CIRS-G scores were generated from chart review, and education level was asked of the participants after administration of study measures.

Therapists administered 6 of 7 tasks of the CPT. The therapists were not able to feasibly administer the “phone” task due to lack of a working phone for patient use which is acceptable since the test manual states that one or two tasks may be eliminated from the test. The Assessment of Motor and Process Skills (Fischer, 2006) was administered by an outside assessor who was blind to the score on the CPT. To ensure that completion of the AMPS followed the assessment protocol, the AMPS administrator was aware of the patient medical history and SMMSE score. The AMPS version 7 cut-point scores were

used to indicate independence (Merritt, 2010).

The CPT and SMMSE, as part of usual care, were administered within several days of admission to the rehabilitation unit. Due to the consent process, and feasibility of administration by the outside assessor, time elapsed between the administration of the usual care measures (CPT, SMMSE and FIM) and the AMPS scales. The average number of days between administration of the CPT and AMPS was 4.7 days (SD= 2.7 days).

Analyses

Internal consistency reliability of the CPT was determined using Cronbach's alpha. Determination of the influence of age, sex and education was conducted using linear regression with CPT score as the dependent variable and age, sex and education as independent variables. Discriminant validity of the CPT compared with measures of ADL, chronic medical illness, motor skills, was analyzed using Pearson's r correlations. Concurrent validity with cognitive screening and process skills was also analyzed with Pearson's r correlations. For the final objective, ROC curves were calculated to examine the sensitivity and specificity of the CPT to predict "independence" according to the SMMSE and AMPS scale. The ROC curve was used to determine the cut-point which maximized both sensitivity and specificity of the CPT. Using this cut-point, the strength of agreement between the CPT and both measures (SMMSE and AMPS-Process scale) for designation of "independence" was determined using the phi statistic.

Results

Table 1 describes the sample demographics. The sample (n=47) was composed of a majority of women (55.3%). Participants had a mean age of 83.5 years with 14 participants ≥ 85 years. A majority of participants (87.2%) had either grade school or high school education with a minority (12.8%) having post-secondary education.

The internal consistency of the CPT was acceptable (Cronbach's $\alpha=0.71$). The subtest-total correlation was acceptable at $r > 0.2$ for each subtest total score compared to the overall test score. Subtest-total statistics demonstrated that no subtest acted uniquely relative to any another. The ability of any single subtest to predict the scores on other subtests was low (ICC=0.3).

Table 1: Description of sample

	Valid (N)	Range	Mean	SD
Gender				
male	21 (44.7%)			
female	26 (55.3%)			
Age		66 – 97	83.5	7.7
Years of Education		2 – 17	10.2	3.1
FIM score		76 -119	103.09	12.25
CIRS-G		1.30 - 2.60	2.04	0.27
SMMSE score		10 – 30	23.79	4.05
AMPS-Motor		-0.25 - 2.24	0.95	0.54
AMPS- Process		-0.42 -1.49	0.66	0.45
CPT score		3.8 - 5.5	4.53	0.38

CPT scores were not significantly influenced by age, sex or years of education.

This was shown by linear regression with CPT as the dependent variable and age,

sex and years of education as independent variables. This model had an overall F-test that was not significant.

Correlations between scores on the CPT and validation measures revealed the anticipated pattern (see Table 2). When testing discriminant validity, a non-significant correlation was found with measures of motor skills (AMPS-Motor: $r=0.15$) and chronic medical illness (CIRS-G: $r=-0.26$). There was a weak correlation with the measure of ADL burden (FIM: $r=0.32$, $p<0.05$). When testing concurrent validity, significant correlations were found with measures of cognition (SMMSE: $r=0.47$, $p<0.01$) (AMPS-Process: $r=0.53$, $p<0.01$).

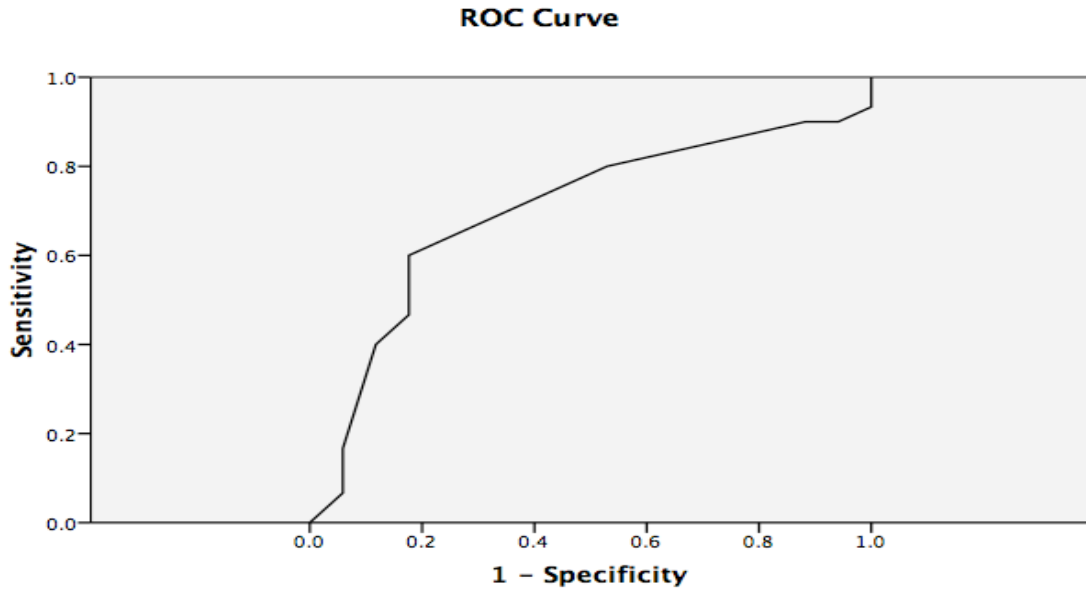
Table 2: Pearson correlation coefficients of CPT with validation measures

Measure	FIM	CIRS-G	AMPS-Motor	SMMSE	AMPS-Process
Cognitive Performance Test (CPT)	0.32*	-0.26	0.15	0.47 **	0.53**
Functional Independence Measure (FIM)		-0.09	0.62**	0.19	0.33*
Cumulative Illness Rating Scale- Geriatric (CIRS-G)			-0.28	-0.38**	-0.22
AMPS-Motor				0.11	0.43**
Standardized Mini Mental Status Exam (SMMSE)					0.46**

* significant at p<.05 level ** significant at p<0.01 level

The ability of the CPT to predict a designation of impairment on the SMMSE and AMPS was examined using ROC curves. The areas under the ROC curve were similar when comparing the CPT to the SMMSE (Area under curve= 0.70) and AMPS-Process scale (Area under curve= 0.79) (Figures 1 and 2). The sensitivity and specificity of the CPT for predicting impairment according to the SMMSE was optimized at a CPT cut-point of 4.5. At this point the sensitivity of the CPT was 0.53 and specificity is 0.82. The cut point that optimized sensitivity and specificity related to the AMPS was a CPT score of 4.6. At this point the sensitivity was 0.75 and specificity was 0.65.

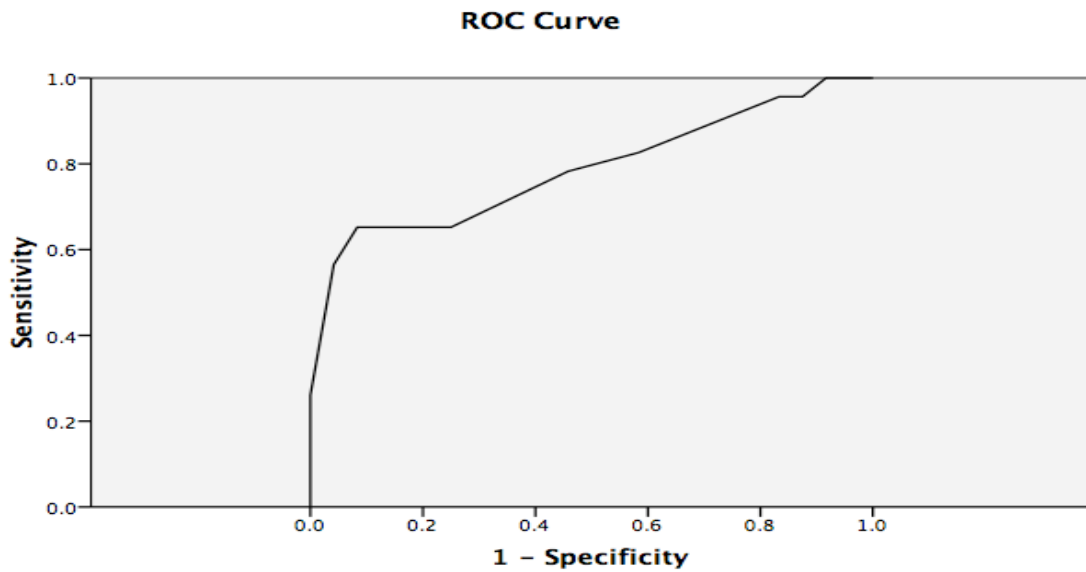
Figure 1: ROC curve for CPT with SMMSE as criterion



Diagonal segments are produced by ties.

Area under curve= 0.70

Figure 2: ROC curve for CPT with AMPS-Process scale as criterion



Diagonal segments are produced by ties.

Area under curve =0.79

Using the CPT cut score generated from the ROC curve, the overall agreement between the CPT and two criterion measures (SMMSE and AMPS-Process) was examined using the phi statistic. Results are shown in tables 3 and 4. The strength of the association (phi statistic) was statistically significant when comparing the CPT to the SMMSE and AMPS-Process scale.

Table 3: Phi test: CPT score and SMMSE cut-off

CPT total score	SMMSE designation		Total
	0-26 Impaired	>26 (Unimpaired)	
< 4.5 (impaired)	16	3	19
4.5 or higher (unimpaired)	14	14	28
Total	30	17	47

phi= 0.35 p=0.02

Table 4: Phi test: CPT score and AMPS-P cut-off

CPT total score	AMPS-Process designation		Total
	Dependent	Independent	
< 4.6 (impaired)	18	8	26
4.6 or more (unimpaired)	6	15	21
Total	24	23	47

phi= 0.40 p=0.01

Discussion

Internal consistency

The internal consistency of the CPT was acceptable, indicating that subtest scores were reasonably well correlated with one another. The single measures intra-

class correlation coefficient indicated that one could not reliably predict another subtest score from any single subtest and that multiple subtests need to be administered in order to obtain acceptable internal consistency. Therapists administering the CPT need to consider that there is a cost associated with reducing the number of subtests for clinical feasibility, in that the reliability of the overall score is reduced. Whenever possible, to maximize reliability, therapists should administer all of the CPT subtests.

Lack of effect of age, sex and number of years of education

In this study, age, sex and number of years of education did not significantly relate to scores on the CPT. The scores did not reflect educational attainment, nor were they affected by declines in age-related factors such as sensory decline or mental speed of processing. Additionally, it is valuable for a measure of cognition in daily living to be minimally influenced by sex and educational attainment. The score will more accurately reflect a person's ability to perform tasks on a daily basis (regardless of gender) than the ability to successfully negotiate an assessment or testing environment. If a measure is influenced by education, persons with lower levels of education may be unfairly disadvantaged. These data indicate that the CPT was performing in this sample independently of age, sex and educational attainment.

Lack of correlation with motor skills and co-morbidities

CPT scores did not correlate with a measure of motor skills (AMPS-Motor scale) or co-morbidities (CIRS-G) (Table 2). The CPT was designed as a measure of cognition and these findings provide evidence of discriminant validity for the CPT,

indicating that it is not principally measuring motor skills or physical factors related to diagnosis. The subtests of the CPT include daily living tasks such as washing hands and making toast. These require physical motor skills such as postural control, grasp, and hand coordination, but any influence that physical motor skills may have had on the scores of the CPT were not observable in this sample. The instructions for CPT administration specify ways to minimize the effects of motor ability on the overall score of CPT by, for example, moving the toaster closer to the person if it is too difficult to reach. The author of the CPT intended to reduce the effect of motor skills on scoring, and the data from this study indicate that, indeed, any effect of motor skills on the scoring is minimal.

The correlation with ADL burden of care (FIM) was weaker than the correlation with cognitive screening or process measures but stronger than with measures of motor skills and co-morbidities. This can be explained because FIM scores are influenced by both cognitive and motor skills. The FIM describes independence in tasks such as dressing and washing, such that those with a lower level of cognition require assistance and score lower on the FIM. Compared with measures of motor skill alone, FIM scores should therefore, be more highly correlated with CPT. However, compared with measures of cognition alone, FIM scores are influenced by motor skill, this weakening the correlation with CPT.

Correlation with cognitive screen and process skills

The validity of the CPT was supported by its correlation with measures of cognitive screening (SMMSE), and process skills (AMPS Process scale), thereby

indicating that each are measuring a similar construct. The finding of moderate correlation between the CPT and SMMSE replicates previous findings of correlation with the MMSE (Burns, 1992; Bar-Yosef et al., 1999). Each are designed to measure the overall construct of cognition, but the moderate correlation indicates either both are associated with error or measure cognition differently. Because the reliability of both measures is reported to be high (CPT $r > 0.8$ (Burns, 1992); SMMSE $r > 0.8$ (Molloy, Alemayahu & Roberts, 1991)) the correlation when corrected for reliability remains moderate. The SMMSE is designed to measure cognitive components including memory, attention, and abstraction whereas the CPT examines working memory using daily living tasks. Furthermore, the SMMSE is designed as a screening tool, and the CPT is designed to give information about the level of cueing needed to do a task with the aim of treatment planning. The differences between the tools in overall design and purpose can explain the moderate correlation between scores.

The statistically significant correlation with the AMPS-Process scale provides additional evidence that the CPT is a measure of cognition. The AMPS-Process scale is correlated with measures of cognition including the LOTCA, Cognistat and Rey Complex Figure-copy (Kizony & Katz, 2002). Correlation with the AMPS-Process scale adds to previous evidence suggesting a correlation between CPT and neuropsychological measures of planning, sequencing, and attention (Bares, 1998). The finding of a moderate correlation suggests that the CPT and AMPS-Process are unique in their measurement of cognition. A primary difference may lie in the

theoretical construction and scoring of the measures. The CPT is designed to measure working memory and scoring is based on the amount of cueing required to complete tasks (Burns, 1992). The AMPS is described by Fisher (2006) to be a measure of observable behaviours, called “process skills” rather than cognitive components such as memory or attention. Therefore scoring for the AMPS is based on observable actions, such as placement of items in the work space, and task sequence. Another difference that may affect measurement is that the same CPT tasks are administered consistently, whereas the AMPS tasks are chosen by the patient. Although both use everyday tasks, the act of choosing a task in the AMPS may increase attention to the task, care in task completion, and perseverance in the face of difficulty.

Designation of impairment compared to other measures

The ability of the CPT to predict whether a person would score “impaired” on the CPT was examined using the cut-off points for the SMMSE and AMPS-Process scale. ROC curves were generated for the CPT versus SMMSE and AMPS-Process scales (Figures 1 and 2). The area under the ROC curve was similar when comparing the CPT to AMPS-Process scale and the CPT to SMMSE. This indicates that agreement between CPT scores and the designation of “independent” on the AMPS-Process scale was similar to the agreement with the SMMSE.

The CPT was similar in designation of impairment compared to both criterion measures (SMMSE and AMPS-P) according to the significant phi coefficients. This finding differs from data comparing a measure similar to the CPT with the AMPS-

Process scale (Marom, Jarus & Josman, 2006). Scoring of the CPT is based on the cognitive levels described in the test manual for the Large Allen Cognitive Levels Test (LACL) (Marom, Jarus & Josman, 2006). Therefore the difference in findings is important to note. In the previous study, no relation was found between the LACL and AMPS-process cut-off scores. The authors concluded that the LACL and AMPS differed in constructs being assessed. The finding of a relationship between CPT and AMPS in this study may indicate that the constructs assessed do not differ as much as previously concluded. The finding of a relationship between CPT and AMPS in this study may be due to use of a cut-point using ROC curves, whereas the previous study used three separate groups for CPT and two groups for AMPS in the analysis. Both studies reported similar sample demographics and range in scores. This current study was able to demonstrate an association with AMPS and therefore provide construct validation evidence for the CPT.

In this study, although the CPT performed similarly to the SMMSE and AMPS-Process scales in distinguishing impaired from unimpaired persons, the strength of the association was weak (phi coefficient less than 0.5). The degree of agreement between the CPT and SMMSE was higher for those who were not impaired. That is, persons who scored poorly on the CPT were more likely to score impaired on the SMMSE or AMPS-Process Scale. However, the sensitivity and specificity for detection of impairment was not consistent in all cases in that some persons who scored impaired on one measure were unimpaired on the other and vice versa. The

conservative conclusion is that there was error associated with measurement in both instruments that resulted in differing designations of impairment.

When considering the use of the CPT in clinical practice, it is important to examine the number of persons who are designated as impaired by the measure. This informs the user about whether one measure was consistently more stringent at designating impairment than another. In this study, the CPT was not consistent with the SMMSE or AMPS-Process scale in designating impairment. This means that whether a person is identified as having impairment may depend on the measure chosen by the therapist, but also on error associated with measurement. The pattern of designation of impairment was not consistent in all cases. Clinical interpretation of the results must take into account the degree of uncertainty in identifying impairment. This study provides data that supports the construct validation of the CPT and test scores must be interpreted alongside other clinical observation of cognition and function when working with people with dementia.

Limitations and directions for future research:

A higher alpha level for internal consistency may have been obtained if all subtests had been administered, however this was not feasible in the clinical setting and the alpha level obtained was acceptable. The CPT test was administered in the course of usual care thus therapist administrator and time may be confounding the correlation between measures. Two therapists administered the CPT, and although they both had training given by the CPT author, inter-rater reliability was not established and may have reduced the strength of the correlations found. Also,

clinical feasibility meant that there was a time lag of days (average 4.7 days) between measures. This time lag was similar to a previous study (Kizony & Katz, 2002), the participants were medically stable, and they did not display symptoms of delirium. It is likely that the participants would not have tolerated administration of all measures on a single day, and fatigue could negatively influence test performance. The correlations were significant as hypothesized, but they may have been stronger if the measures had been administered on the same day or closer together in time.

The range of CPT scores in the sample was restricted from 3.8 to 5.5. Restriction in range with a high number of persons at the middle range of the scores places a limit on finding a strong correlation (Streiner & Norman, 2003). The sample was drawn from an inpatient rehabilitation unit and although potentially representative of geriatric inpatients, was not representative of community-dwelling older adults. Further restriction in range was avoided by not requiring the sample to have a diagnosis of dementia, merely a suspicion of cognitive deficit, and persons who were not suspected to have cognitive deficits were not on the caseload of the clinicians administering the test measures. Correlations may have been stronger if there had been a greater range in CPT scores.

Further research is required to assess the predictive validity of the CPT to determine safety at home upon discharge from rehabilitation. Occupational therapists need “top-down” performance-based measures for predicting function and making recommendations for discharge from rehabilitation. There is a call for

cognitive tests that are ecologically relevant, and based on real-world performance, as they will provide an empirical foundation for decisions on safety (Burgess, Alderman, Forbes, Costello, Coates, Dawson, et al. 2006). The CPT, as a performance-based test is well situated to meet this need. Moreover, because the CPT examines both cognitive capacities and daily function, it can be used for both identifying deficits and predicting function. Therefore, it is attractive as an “efficient” measure (Milberg, 1996). This study demonstrates that the CPT is a valid measure of cognition that was not influenced by age, sex, or education or motor function. The interpretation of the score must be accompanied by a statement regarding evidence for error in identifying impairment in a sample of older adults with cognitive impairment. Further development of occupational therapy assessments of function will promote ethical decision-making in the care of persons with dementia.

References

- American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders* (text revision). Washington, D.C.: Author.
- Bares, K. K. (1998). Neuropsychological predictors of functional level in Alzheimer's disease. Ph.D. dissertation, The Herman M. Finch University of Health Sciences - The Chicago Medical School, Illinois, United States. Retrieved February 19, 2010, from: *Dissertations & Theses: A&I*. (Publication No. AAT 9945838).
- Bar-Yosef, C., Weinblatt, N., & Katz, N. (1999). Reliability and validity of the Cognitive Performance Test (CPT) in an elderly population in Israel. *Physical and Occupational Therapy in Geriatrics*, 17, 65-79.
- Bonder, B.R. & Wagner, M.B. (2001). *Functional Performance in Older Adults* (2nd ed.) Philadelphia: F.A. Davis Company.
- Burns, T. (1992). The Cognitive Performance Test: An approach to cognitive level assessment in Alzheimer disease. In C.K. Allen, C.A. Earhart & T. Blue (Eds.), *Occupational therapy treatment goals for the physically and cognitively disabled* (pp. 46–84). Rockville, MD: The American Occupational Therapy Association, Inc.
- Burns T., Mortimer J.A., & Merchak, P. (1994). Cognitive Performance Test: A new approach to functional assessment in Alzheimer's disease. *Journal of Geriatric Psychiatry and Neurology*, 7, 46-54.
- Burns, T. (2006). *Cognitive Performance Test Manual*. New Jersey: Maddak.
- Burgess, P. W., Alderman, N., Forbes, C., Costello, A., Coates, L., Dawson, D. R., et al. (2006). The case for the development and use of “ecologically valid” measures of executive function in experimental and clinical neuropsychology. *Journal of the International Neuropsychological Society*, 12, 194–209.
- Crepeau E.B., Cohn, E.S., & Schell, B.B. (2003) Occupational therapy practice. In E. B. Crepeau, E. S. Cohn, & B., Boyt-Schell. (Eds.) *Willard & Spackman's Occupational Therapy* (10th edition, pp 27-30). Philadelphia: Lippincott, Williams & Wilkins.
- Duchek, J.M. & Abreu, B. (1997). Meeting the challenges of cognitive disabilities. In C. Christianson, & C. Baum, (Eds.). *Occupational Therapy: Enabling Function and Wellbeing 2nd edition*. New Jersey: Slack Inc.
- Fisher, A. G. (2006). *Assessment of Motor and Process Skills. Vol. 2: User Manual* (6th

ed.) Fort Collins, CO: Three Star Press.

- Folstein, M.F., Folstein, S.E., & McHugh, P.R. (1975). "Mini-Mental State": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12, 189-198.
- Gaugler, J.E., Duval, S., Anderson, K.A. & Kane, R.L. (2007). Predicting nursing home admission in the U.S: a meta-analysis. *BMC Geriatrics*, 7, 961-973.
- Glenny, C. & Stolee, P. (2009). Comparing the Functional Independence Measure and the Inter-RAI/MDS for use in the functional assessment of older adults: a review of the literature. *BMC Geriatrics* 9, 52 doi: 10.1186/1471-2318-9-52
- Grieve, J. (2000). *Neuropsychology for Occupational Therapists*. Oxford: Blackwell Science Ltd.
- Kizony, R. & Katz, N. (2002). Relationships between cognitive abilities and the process scale and skills of the assessment of motor and process skills (AMPS) in patients with stroke. *Occupational Therapy Journal of Research*, 22, 82-92,
- Law, M. Baum C. & Dunn, W. (2005). *Measuring Occupational Performance: supporting best practice in occupational therapy* (2nd ed.). Thorofare, NJ: Slack.
- Lawton, M.P. (1971). The functional assessment of elderly people. *Journal of the American Geriatrics Society*, 19, 465- 481.
- Letts, L., & Bosch, J. (2005). Measuring occupational performance in basic activities of daily living. In: M. Law, C. Baum & W. Dunn (Eds.). *Measuring Occupational Performance: supporting best practice in occupational therapy* (2nd ed.). Thorofare, NJ: Slack.
- Marom, B., Jarus, T. & Josman, N. (2006). The relationship between the Assessment of Motor and Process Skills (AMPS) and the Large Allen Cognitive Level (LACL) Test in clients with stroke. *Physical and Occupational Therapy in Geriatrics* 24(4), 33-50.
- Merritt, B. K. (2010). Utilizing AMPS ability measures to predict level of community dependence. *Scandinavian Journal of Occupational Therapy*, 17, 70-76.
- McDowell, I., & Newell, C. (1996). *Measuring Health: a guide to rating scales and questionnaires* (2nd ed.). Oxford University Press.

- Milberg, W.M. (1996). Issues in the assessment of cognitive function in dementia. *Brain and Cognition* 31, 114-132.
- Miller, M.D., Paradis, C.F., Houck, P.R., Mazumdar, S., Stack, J.A., Rifai, A.H., Mulsant, B., & Reynolds III, C.F. (1992). Rating chronic medical illness burden in geropsychiatric practice and research: application of the Cumulative Illness Rating Scale. *Psychiatry Research*, 41, 237–248.
- Molloy, D.W. & Clarnette, R. (1999). *Standardized Mini Mental Status Exam (SMMSE): A user's guide*. Troy, Ontario: New Grange Press.
- Molloy, D. W., Alemayehu, E., & Roberts, R. (1991). Reliability of a standardized Mini-Mental State Examination compared with the traditional Mini-Mental State Examination. *American Journal of Psychiatry*, 148, 102-105.
- Molloy D.W., Silberfeld, M., Darzins, P., Guyatt, G.H., Singer, P.A., Rush, B., et al. (1996) Measuring capacity to complete and advanced directives. *Journal of the American Geriatric Society*, 44, 660-664.
- Streiner D.L & Norman, G.R. (2003). *Health measurement scales: A practical guide to their development and use* (3rd ed.). Oxford, UK: Oxford University Press
- Thralow, J.U. & Reuter, M.S. (1993). Activities of daily living and cognitive levels of function in dementia. *American Journal of Alzheimer's Care and Related Disorders & Research*, 8, 14-19.
- Townsend, E.A., & Polatajko, H. (2007). *Enabling occupation II: Advancing an occupational therapy vision of health, well-being and justice through occupation*. Ottawa, ON: CAOT Publications ACE.
- Uniform Data System for Medical Rehabilitation: Functional Independence Measure. Retrieved July 5, 2010 from: <http://www.udsmr.org/>
- Vining Radomski, M. (2002). Assessing abilities and capacities: cognition. In C.A. Trombly & M. Vining-Radomski (Eds.), *Occupational Therapy for Physical Dysfunction 5th edition*. (pp.199-212). Baltimore: Lippincott Williams & Wilkins.
- World Health Organization (2002). Towards a Common Language for Functioning, Disability and Health: ICF The International Classification of Functioning, Disability and Health Retrieved Feb 19, 2010 from: <http://www.who.int/classifications/icf/icfaptraining/en/index.html>

CHAPTER 3:

**Measurement of harm outcomes in older adults after hospital discharge:
Reliability and Validity**

Authors

Alison Douglas
Lori Letts
Kevin Eva
Julie Richardson

Target journal: Alzheimer's Disease and Associated Disorders

**Measurement of harm outcomes in older adults after hospital discharge:
reliability and validity**

Abstract: (200 words)

Purpose: Clinicians use assessment measures to aid in determining if a person with dementia is safe, but measures need validation using safety outcomes. Defining and validating a method for measurement of safety contributes to further validation of clinical measures. The study objectives were to determine the test-retest reliability and validity of measuring the outcome “incidents of harm” through caregiver interviews. Methods: The Incident of Harm Caregiver Questionnaire was designed based on previous research and administered by telephone once per month for six months. Caregivers completed daily logs for one month and medical charts were examined for validation. Results: Test-retest reliability measured over a one-week period (n=38), was high for the occurrence of an incident of harm (yes/no; kappa= 1.0) and the type of incident (percent agreement=100%). Validation against daily logs found no disagreement regarding occurrence or types of incidents. Validation with medical charts found no disagreement regarding the occurrence of an incident of harm and disagreement in half of the six incidents for incident type. Conclusion: The data support the validity of caregiver interview for determination of occurrence of incidents of harm and are important to researchers as a method to measure safety when validating clinical measures.

Key words (3-6): older adult, dementia, accidents, validation

Introduction

Decisions about whether an older adult can live independently require information about safety. Assessment measures currently used by clinicians to aid in determining if a person is safe are typically validated for outcomes that do not include safety. Safety, or the ability to live without unintentional harm or injury, is clinically important in geriatric rehabilitation because of the risk of self-neglect or disorientation, whether or not a person has been diagnosed with frailty or dementia. Concerns arise about unintentional falls, harm from medication errors, or failure to take care of daily needs. Research into safety outcomes is urgently needed to provide information to determine when older adults are at risk of declining health due to self-neglect (Dyer, Pickens, & Burnett, 2007).

Little is known about the ability of assessment tools to predict safety as an outcome. Measures for use with people with dementia have been validated by examining their scores in relation to outcomes such as daily living skills (Cameron et al. 2010; Ruchinskas & Curyto, 2003; Ruchinskas, Singer & Repetz, 2000; Tombaugh, & McIntyre, 1992), discharge status (MacNeill, Gerskovich, Caron, Lichtenberg, 1997; Nadler et al. 1993), community functioning (Velligan et al. 1998) or likelihood of independent living (Ruchinskas & Curyto, 2003; Tombaugh, & McIntyre, 1992). These data, although important for establishing the validity of a tool for measuring real-life outcomes, lack a direct link with predicting safety. It is important to establish a reliable and valid outcome measure that addresses safety in the community for older adults.

Initial work has been done to operationalize safety as “incidents of harm” (Tierney, et al., 2004). Incidents of harm were defined as unintentional events that through self-neglect or disorientation cause: physical injury to self (including exacerbation of chronic illness) or other; and/or damage or loss of property (Tierney et al., 2004). In their study, the number of incidents was assessed by administering a telephone questionnaire to the caregiver. Types of incidents were recorded and classified into the following categories: failure to eat and drink; failure to report a medical condition; failure to use prescribed assistive devices; failure to maintain personal hygiene; failure to take medication properly; failure to recognize a familiar environment; failure to turn off electrical appliances; and failure to judge fraudulent activities. Results of the study demonstrated a link between incidents of harm and scores on cognitive screening measures; however data regarding the consistency or accuracy of caregiver reports was not provided. Further validation of the measurement of “incidents of harm” is needed. The accuracy of caregiver report may be challenged, and data demonstrating its accuracy would allow more definitive conclusions regarding the predictive validity of measures for this outcome.

A telephone interview of a knowledgeable caregiver is a feasible administration method for measuring the outcome “incidents of harm” for several reasons. First, the use of a telephone interview, compared to in-person home visits, requires less time from the study participants, is less burdensome for busy caregivers, and is more cost effective for the research team. Telephone information is likely to be similar to home visit information because in both cases, the research

team must rely on reporting of incidents that have occurred in the recent past.

Although a home visit may afford the ability to examine the physical environment, information about the occurrence of events such as failure to eat or falls can only be garnered from reports by the study participants.

Second, proxy reporting has the advantage of minimizing reporting bias (Streiner & Norman) on the part of the older adult, especially if cognitive impairment is suspected. In a population of patients on a geriatric rehabilitation unit such as in this study, some persons may have diagnosis of dementia, whereas others may have physical deficits only, or suspicion of cognitive deficits. Persons with confirmed or possible dementia are likely to have difficulty recalling recent events, may lack insight into abilities, or may have fear of reporting injury. Proxy reporting may minimize reporting bias also by facilitating the person with dementia, who may be suspicious or lack the ability to understand study information, to have comfort with the research study. A caregiver may be more likely to understand explanations about confidentiality and lack of negative consequences in reporting incidents. Cooperation with the research team by a trusted caregiver may be less likely to prompt suspicion on the part of the person with dementia.

Lastly, the use of proxy reporting by a caregiver is supported by literature reviews of patient-proxy reliability, especially when the report is related to observable constructs. Accumulated evidence supports that scores obtained from patients and proxies have good agreement. Patient-proxy agreement is highest for measures that examine observable behaviours such as eating, and dressing

(Oczkowski & O'Donnell, 2010) or level of functioning (Neumann, Araki & Gutterman, 2000). Less observable outcomes such as quality of life or satisfaction showed lower levels of agreement (Oczkowski & O'Donnell, 2010; Neumann, Araki & Gutterman, 2000).

The objectives of the study were two-fold: first, to determine the test-retest reliability of measuring the outcome “incidents of harm” by caregiver interviews and second, to determine the validity of measuring the outcome “incidents of harm” compared to a) physician reports on medical charts and b) caregiver daily logs. The Incident of Harm Caregiver Questionnaire was designed based on the definition of “incidents of harm” from Tierney et al. (2004) (see Appendix 1).

Methods

Sample

Participants were recruited from consecutive admissions to a geriatric rehabilitation unit. Participants were recruited as part of a broader prospective study, the purpose of which was to examine the predictive validity of pre-discharge measures to predict post-discharge “incidents of harm” for six months after discharge. The local research ethics board approved the study procedures. Consent was obtained for the research team to access medical records, for the caregiver to be contacted for a telephone interview once per month for six months after the patient was discharged, and for the family physician to be contacted if there was any change in health status. Signed consent from both the patient and caregiver was required.

Patient and caregiver pairs were eligible for the study if the patient was over age 65, undergoing assessment for possible dementia (diagnosis of dementia was not required), and the caregiver was identified by the patient and caregiver as the most knowledgeable person about the patient. Exclusion criteria were delirium or a query from the health care team of elder abuse.

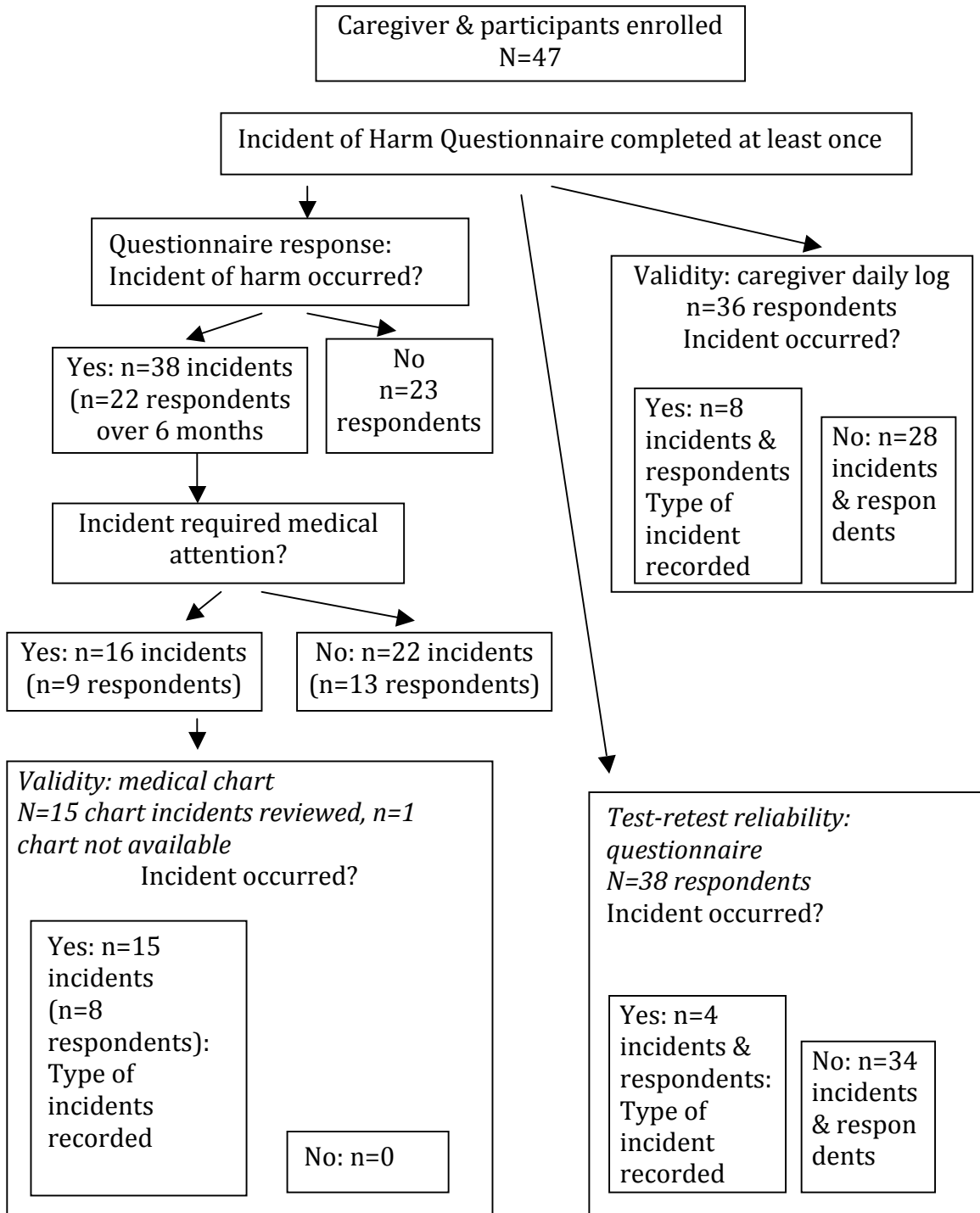
Procedures

For this study, the data collection described was completed after patients were discharged from hospital during the six month follow-up period. A research assistant, who was blind to medical record information, contacted the caregivers by telephone and administered the Incident of Harm Caregiver Questionnaire once each month for six months (see Appendix 1 for questionnaire). The questionnaire asked the caregiver to identify whether any incident of harm had occurred (yes/no) and whether it was due to self-neglect or disorientation (yes/no) (see Figure 1 for study procedures). If an incident was reported, the caregiver was requested to identify from a list the most likely reason for the occurrence (type of incident). If no incident of harm was noted, the interviewer asked, “So is everything going OK?” to gain information about overall functional status, and give further opportunity to ask about possible incidents of harm.

For test-retest reliability, caregivers were contacted up to five days after one monthly interview and re-administered the questionnaire. The interviewer attempted to contact the caregiver for test-retest administration after the third monthly interview, but if contact was not made within five days, the test-retest

administration occurred in a later month. The number and types of incidents of harm were compared between the two interviews.

Figure 1: Study procedures



For validity, medical record information was reviewed to give a comparison between caregiver and physician report for any incident that required medical intervention. If information about an incident requiring medical attention was not available on the electronic chart, the family physician was contacted by fax to complete a one page questionnaire (Tierney, et al., 2004, adapted with permission from Dr. Mary Tierney) Caregivers were also requested to keep a daily log of number and types of incidents for either the first or second month after discharge. These were returned by post to the researcher. If the caregiver failed to return the daily log, the researcher phoned to request that it be completed for the current month. New forms were mailed or hand-delivered if requested by the caregiver.

Analyses

Test-retest reliability for incident occurrence (yes/no) was calculated using kappa, and for type of incident (failure to complete ADL, etc.) using proportion agreement. Data were collected on the total number of incidents of harm reported by caregivers and the total number of these requiring medical interventions. Only the incidents that required medical intervention were verifiable by checking medical information. The validity of caregiver reported incidents of harm was tested against physician reports and caregiver daily logs using kappa for incident occurrence (yes/no) and proportion agreement for type of incident.

Results

The sample of patients (n=47) was composed of 55% females with a mean age of 83.3 years including 14 participants over the age of 85. A total of 45 persons contributed data to the study and two participants dropped out by declining to answer or return follow-up calls. Of the 45 persons, 39 completed 6 months of follow-up, 2 deceased, 2 caregivers dropped out while consenting for their information to remain in the study, and 2 became ineligible due to admission to long-term care facilities. Responses were grouped into categories according to the method of Tierney (2004) for all incidents (n=35). The highest proportion of incident was of the type “failure to use mobility devices correctly” (n=11, 31.4%), followed by “failure to use medication correctly” (n= 7, 20.0%), “failure to perform ADL” (n=6, 17.1%), “failure to report medical condition” (n=6, 17.1%), and “unknown” type (n=5, 14.3%). Test-retest reliability was obtained from n=38 caregivers, in the fifth month on average (SD= 1.1), with minimum of two and maximum of five days between test and retest (mean= 3.1, SD= 1.1). The test-retest reliability was high for the occurrence of an incident of harm each month (yes/no)(kappa= 1.0) and the type of incident (percent agreement=100%) (Table 1).

Table 1: Test-retest reliability and validity of Incidents of Harm Questionnaire

	Response (N)		N with agreement	Valid N	Analysis	P value
	Time 1	Time 2				
Test- retest reliability						
Incident occurred (yes/no)	yes= 4 no= 34	yes= 4 no= 34	38 38	38 38	Kappa= 1.0 percent agreement= 100%	<0.01
Type of incident						
Validity: medical chart (only incidents with medical visit, n=15)	Caregiver	Medical Chart				
Incident occurred (yes/no)	yes= 15 no= 0	yes= 15 no= 0	15 3	15 6	Kappa= 1.0 percent agreement = 50%	<0.01
Type of incident						
Validity: caregiver daily log	Phone	Log				
Incident occurred (yes/no)	yes=8 no=28	yes= 8 no= 28	36 36	36 36	Kappa= 1.0 percent agreement= 100%	<0.01
Type of incident						

Caregiver reports were validated by checking medical charts or by contacting the family physician regarding any incident that required a medical visit. There were a total of 16 incidents of harm reported by the caregiver that required a medical visit. For a small number of incidents (n=3), information was not available

on the hospital chart and verification was obtained by contacting the family physician. A response was received for two of these three incidents and the third incident was noted as a missing data point because of non-response. There was no disagreement with medical charts ($\kappa=1.0$)($n=15$) for the occurrence of an incident of harm (yes/no)(Table 1).

The type of incident was compared between the caregiver report and notes found on the medical chart ($n=15$) (Table 2). Any information regarding the possible reason for the medical visit was obtained by reading the entire chart, including physician and other health care team members. The medical chart did not report the type of incident (unknown) in the majority of cases ($n=9/15$, 60.0%). The caregiver stated the incident type was unknown in $n=1/15$ (6.7 %). For those incidents for which a type was reported ($n=6$), there was agreement in 3/6 cases (50%). For those that require medical attention, the most frequent type of incident reported by the caregivers was failure to take medication correctly ($n=8/15$), and in the majority of these cases ($n=6/8$, 75%), the medical chart did not report a type of incident.

Table 2: Validity comparison between caregiver and medical chart on type of incident (incidents requiring medical attention only)

Type described by caregiver	Type described by medical chart					
	Failure to do activities of daily living	Failure to take medication correctly	Failure to use mobility devices correctly	Failure to take care of medical condition	Other	Unknown
Failure to do activities of daily living				1		1
Failure to take medication correctly		1		1		6
Failure to use mobility devices correctly			1			
Failure to take care of medical condition				1		
Other (substance use)						1
Unknown			1			1

The validity of the telephone interview method was also examined by asking caregiver participants to track incidents for one month using a daily log. The return rate for daily logs was 36/47 (76.6%). 20/36 respondents needed more than one reminder to complete the log, or additional mailing of a new copy, having lost the first copy. There was no disagreement between the diary and phone calls about whether the incident occurred ($\kappa = 1.0$, $p < 0.001$) and the types of incidents (percent agreement = 100%).

Discussion

The data support the overall reliability and validity of the measurement of harm outcome. First, the test-retest reliability of the caregiver report by telephone was very high. It was apparent that events of harm that were reported were events that were not easily forgotten in a few days. Nor did recall of the type of event change. There was also high agreement between caregiver report and medical chart data. This indicates the caregiver report was valid as a measurement of incident of harm outcome using proxy report. In the cases where there was disagreement between the caregiver and the medical chart, the information from the caregiver was more specific about the potential type of incident compared to the medical chart.

Although proxy reporters may not have been present for all incidents, such as falls or medication errors, the high agreement with medical chart data demonstrates a lack of bias in caregiver reporting when incidents were serious enough to require medical care. This indicates that any incident noted by the caregiver was reported to the research team. The incident of harm outcome can be viewed as a more objective and observable outcome than, for example, satisfaction, depression or optimism which have been noted to be less objective and therefore more difficult for proxies to rate accurately (Oczkowski & O'Donnell, 2010). The

validity of proxy report for the incidents of harm outcome concurs with the literature reporting good patient-proxy agreement with observable behaviours.

Strong agreement between the telephone interview and daily log further validates the data collection method. It indicates that the method of reporting incidents of harm did not affect whether or not an incident was reported. Although the telephone interview did not specifically ask about the number of incidents, the data from the daily logs showed that some caregivers had noted more than one incident in a month. The telephone interview asked for recall of the entire month, suggesting that it may not have been collecting the same level of detailed data as the diary. Future studies may benefit from modifying the telephone questionnaire to gather data on not only whether an incident occurred in the past month, but also how many occurred. However, the high agreement between daily log and telephone interview for reporting whether an incident occurred validates the telephone interview as an appropriate method for collection of data regarding the occurrence of incidents of harm. Furthermore, the return rate, despite multiple reminders and effort on the part of the researchers indicated that the daily log was more burdensome for the caregivers, thus making the telephone interview the preferred method of data collection.

The data from this study indicate that the measurement of incidents of harm through telephone interviews can provide a valid estimate of whether or not an incident occurred according to caregivers and medical staff for this sample. This measurement method is important for validation of clinical assessment tools. It can

be used to validate tools for outcomes related directly to safety in the home environment. The outcome “incidents of harm” is a “real-life” outcome that has been noted to be of urgent need in geriatrics (Dyer, Pickens, & Burnett, 2007) and clinical assessment (Baird, Podell, Lovell, & Bell-McGinty, 2001). The outcome “incidents of harm” has value for validation of health care measurement tools used in a broad range of populations and health care settings. The ability to predict community safety is important for clinical assessment with populations such as people with mental illness and acquired brain injury, and the “incidents of harm” outcome quantifies the concept of safety in the real-world setting. The outcome is particularly pertinent in the older adult population for which decisions about independent living and the need for community services often depend on the determination of safety.

Limitations

Several sources of bias may have influenced the results. The caregivers may have been influenced by reporting and recall bias. Caregivers assume a burden by caregiving, which adds to the stress from various other roles including work and parenting. Caregivers may therefore have reported incidents of harm based on their perceived tolerance for risk to the individual. Those who were more anxious about risk may have reported more incidents, those less anxious may have reported an incident but may have stated its consequences were minor and no significant injury occurred. Thus, recall and reporting bias may have either raised or lowered the number of incidents reported. However, when reports by the caregivers were

checked against medical charts, the agreement was high, indicating minimal reporting bias for incidents requiring medical intervention.

Reliability data were based upon reports from telephone interviews of the caregiver, therefore are limited because both sources of data come from the same person. Validity data, although checked against medical records, were also validated using daily logs from the same person, through a different reporting method.

Proxy report may have lowered the number of incidents reported because proxies may not have been aware of all incidents. However, as previously described, proxy report may have minimized the likelihood of reporting bias on the part of the older adult. The medical chart also may have been influenced by recall and reporting bias such that the medical team may have suspected that the medical visit had been precipitated by disorientation or self-neglect but may not have noted any suspicions on the chart.

Conclusions

The telephone interview method for measurement of incidents of harm demonstrated good test-retest reliability and validity in this sample. The measurement of “incidents of harm” by telephone interview with a proxy caregiver showed high agreement with daily logs, while having higher return rates than daily logs. Moreover, it was validated against medical chart reports. Conservative conclusions are that the use of the data from this outcome measurement can be valid as an estimation of the true occurrence of incidents of harm. More broadly, these data can be used to support the further use of this method of measurement of

incidents of harm in research, especially because the outcome “incidents of harm” reflects how a person is managing in the community, for which there is an urgent need.

References

- Baird, A., Podell, K., Lovell, M., Bell-McGinty, S. (2001). Complex real-world functioning and neuropsychological test performance in older adults. *The Clinical Neuropsychologist* 15 (3): 369-379.
- Cameron J., Worrall- Carter, L., Page, K., Riegel, B., Kai Lo, S., Stewart, S. (2010). Does cognitive impairment predict poor self-care in patients with heart failure? *European Journal of Heart failure* 12, 508-515.
- Dyer, C. B., Pickens, S., & Burnett, J. (2007). Vulnerable elders: When it is no longer safe to live alone. *Journal of the American Medical Association*, 298(12), 1448-1450.
- MacNeill, S. E., Gerskovich, T., Caron, J., & Lichtenberg, P. A. (1997). Living alone: Predictors of recovery during medical rehabilitation. *Clinical Gerontologist*, 18(1), 3-13.
- Neumann, P.J., Araki, S.S., & Gutterman, E.M. (2000). The use of proxy respondents in studies of older adults: lessons, challenges and opportunities. *Journal of the American Geriatric Society*, 48, 1646-1654.
- Oczkowski, C. & O'Donnell, M. (2010). Reliability of proxy respondents for patients with stroke: a systematic review. *Journal of Stroke and Cerebrovascular Disease* 19, 410-416.
- Ruchinkas, R.A. & Curyto, K.J. (2003). Cognitive screening in geriatric rehabilitation. *Rehabilitation Psychology*, 48, 14-22.
- Ruchinkas, R.A., Singer, H. K., Repetz, N. K. (2000). Cognitive status and ambulation in geriatric rehabilitation: Walking without thinking?. *Archives of Physical Medicine & Rehabilitation*, 81, 1224-1228.
- Streiner, D.L., & Norman, G.R. (2003). *Health Measurement Scales: A practical guide to their development and use* (3rd ed.). New York: Oxford University Press.
- Tierney, M. C., Charles, J., Naglie, G., Jaglal, S., Kiss, A., & Fisher, R. H. (2004). Risk factors for harm in cognitively impaired seniors who live alone: A prospective study. *Journal of the American Geriatrics Society*, 52, 1435-1441.
- Tombaugh, T. N., & McIntyre, N. J. (1992). The Mini Mental State Exam: A comprehensive review. *Journal of the American Geriatric Society*, 40, 922-935.

Appendix 1:
Incident of Harm Caregiver Questionnaire
(reference: Tierney et al, 2004)

Follow up interview checklist: harm outcome

1. **Name of patient participant:**

2. **Outcome: Incidents of harm:**

Can you tell me if any of the following have occurred with your___(eg parent, spouse) in the past month?

___(a) any injury to self or another person (including exacerbation of a chronic illness), property loss, or property damage

___(b) occurrence of the incident due to self-neglect or behaviours related to disorientation (unintentional)

3. In your opinion, what might have been the **most important reason** for the occurrence? (Type of incident)

ADL:

Not completing ADL (eating, hygiene, dressing)

Not using medications correctly:

Not using mobility devices correctly:

IADL:

Not using kitchen or electrical devices correctly:

Not detecting fraud:

Not reporting a medical condition:

Environment:

Not recognizing a familiar environment:

Unsafe home set-up:

Other:

4. Note whether medical attention was sought

5. If no incident of harm is reported, ask “So, is everything going OK?”

CHAPTER 4:

Validity of pre-discharge functional measures for predicting harm in older adults

Authors

Alison Douglas
Lori Letts
Julie Richardson
Kevin Eva

Target journal: American Journal of Occupational Therapy

Validity of pre-discharge functional measures for predicting harm in older adults

Abstract

Objective: The purpose of this study was to determine whether two pre-discharge functional measures were valid for predicting incidents of harm after discharge.

Method: Participants (n=47) were recruited from an inpatient rehabilitation unit.

The Assessment of Motor and Process Skills (AMPS) and Cognitive Performance Test (CPT) were administered in the hospital. Independent variables were age, sex, education, living alone, activities of daily living, comorbidities, cognitive screen, and caregiving hours. Incident of harm outcome was measured by caregiver telephone questionnaire monthly for six months.

Results: Compared with all independent variables, AMPS-Process scale was the most significant predictor of harm outcome ($r= 0.37$, $p<0.05$). CPT had a high specificity (91%) for identifying persons who did not have harm. Living alone, age and sex contributed to the prediction of harm.

Conclusion: Both AMPS and CPT demonstrated predictive validity for harm outcome over less predictive variables such as co-morbidities and ADL burden of care.

5 MeSH key words: outcome assessment, psychometrics, older adult, safety, patient discharge

Introduction

In light of increasing demands to maximize bed utilization, health care professionals are pressed to discharge patients as soon as they are medically stable. Discharge planning involves determining the need for home care or long-term care services. Decisions regarding the need for services often rest on assessments of safety and cognition. Older adults with cognitive deficits (e.g. dementias) are at risk because of declining functional skills and the potential for accidental injury. Daily living skills are known to deteriorate as dementias progress (Corcoran, 2001) but the link between these and safety is not well understood. Effective discharge planning for older adults includes the consideration of safety and potential for harm. Therefore, assessment measures that are validated for determining safety are required.

Occupational therapists (OTs) working with older adults with dementia often use non-standardized functional assessments to observe cognitive skills and predict safety (Douglas, Liu, Warren & Hopper, 2007), but previous reviews have concluded that there is little evidence to support their use (Tullis & Nichol, 1999). Many standardized cognitive assessments (e.g. Montreal Cognitive Assessment (Nassredine et al., 2005) are designed to identify impairment rather than predict function, and have not been validated for evaluating safety at home. Although the clinical need to determine safety is common, few standardized assessments used with this population have validity studies linking them with harm outcomes (Douglas, Letts & Liu, 2008). Given the identified gap in evidence for safety (Dyer,

Pickens, & Burnett, 2007), and the potential for adverse outcomes, it is vital that discharge assessments be evaluated.

The risk for physical harm amongst older adults is well known. The leading source of in-home injury for older adults is falls, followed by fires (Lilley, Arie, & Chilvers, 1995). Decreased cognition and polypharmacy are among the greatest risk factors for hospital or emergency admission (Banaszak-Holl et al., 2004). In dementia, studies have identified additional risks including failure to eat, the use of assistive devices and reporting a medical condition (Tierney et al., 2004). Poorer cognition and executive dysfunction are associated with greater risks of nursing home placement (Tun, Murman, Long, Colenda, & von Eye, 2007) and incidents of harm (Tierney, Snow, Charles, Moineddin, & Kiss, 2007).

Validity is the degree of confidence that can be placed in the interpretation of the scores, and as such is a fundamental consideration in the use of tests in health care (Streiner & Norman, 2003). Conceptions of validity have been evolving with the traditional conceptions of validity listing three major types: construct, concurrent and criterion. Subsequently, validity was defined as a single concept, to which multiple types of evidence contribute (Messick, 1995). There are no current standardized guidelines for the evaluation of evidence for health measurement instruments but they are currently being developed by the COSMIN group (Mokkink et al., 2006). Most recently, validation is described as requiring scientific argument using multiple lines of evidence (Kane, 2006). “Argument-based” validity, as Kane (2006) describes it, is seen as an ongoing scientific process. A variety of studies are

required to support the specific use of the assessment. Validation evidence places limits on the interpretation of a test score in that a measure should only be interpreted within the limits of the evidence available. Each intended interpretation of a score must be validated (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999).

In the case of functional measures in dementia, interpretation of test scores to draw conclusions about the risk for injury at home without supporting evidence could lead to problematic decisions about placement in long term care or discharge home. For clinicians to use a measure to predict possible harm outcomes, evidence must demonstrate the relationship between the test score and that particular outcome of harm. Relationship with other outcomes such as disposition to long term care or readmission to hospital contributes less to an accurate interpretation of scores. Interpretation of a test score beyond the bounds of demonstrated relationships to outcomes can lead to unethical use of test scores.

Outcomes currently linked with OT measures include discharge status from rehabilitation or measures of community functioning (Velligan et al. 1998). The Cognitive Performance Test (CPT) and Assessment of Motor and Process Skills (AMPS) show promise for predicting mobility and safety outcomes (Douglas, Letts & Liu, 2008). Low scores on the CPT were associated with a greater risk of institutionalization (Bar-Yosef, et al. 1999). Predictive validity for the AMPS was examined for clients with mental illness (McNulty & Fisher, 2001) and brain injury (Lindén et.al., 2005) using correlation with standardized measures. Data are needed

regarding the ability to predict a clinically relevant harm outcome using a prospective design.

The purpose of this study was to determine the ability of pre-discharge assessment with the Assessment of Motor and Process Skills (AMPS) and Cognitive Performance Test (CPT) to predict frequency of incidents of harm for persons with dementia living at home. The objectives of the study were as follows:

- 1) To determine the ability of CPT score to predict harm outcome, as measured by the incidents of harm reported over a 6-month community follow-up.
- 2) To determine the ability of AMPS- Motor score and AMPS- Process score to predict harm outcome, as measured by incidents of harm reported over a 6 month community follow-up.
- 3) To determine the contribution of independent variables (age, sex, education, comorbidities, activities of daily living independence score, caregiving hours) to prediction of harm outcome.

The design of this study followed the checklist for determination of outcomes, collecting outcomes and follow-up laid out in the STROBE statement for cohort studies (Vandenbroucke, et al., 2007). Independent variables, which could contribute to harm outcomes, were selected based on theoretical constructs relevant to occupational therapy, ease of assessment in occupational therapy, and established relationship to harm outcomes. Therefore co-morbidities (Tierney et al., 2007; Smith, Kokmen, & O'Brien, 2000), activities of daily living (ADL)(Comette et al., 2005) and number of caregiving hours (Smith et al., 2000) were selected.

Methods

The study followed a prospective cohort research design. Participants were recruited from consecutive admissions to a geriatric rehabilitation unit in an urban hospital. The inpatient program is for older adult patients who are medically stable and require short rehabilitation (2-4 weeks). Participants were included if they met the following criteria: 1) 65 years or older, 2) referred to OT, 3) English speaking, 4) undergoing functional or cognitive assessment with the OT (diagnosis of dementia not required to reduce restriction in range of scores on outcome measures) (Streiner & Norman, 2003), and 4) have an identified caregiver for follow up. To reduce recall bias, follow-up data were obtained monthly through a caregiver informant identified by the participant. The caregiver informant had at least weekly in-person contact with the participant to increase accuracy. Exclusion criteria were the following: 1) active delirium, 2) primary diagnosis of mental illness excluding dementia noted on the chart 3) team queried elder abuse 4) planned discharge to a long term care facility.

Procedures

Participants who met the inclusion criteria were identified through review of patient census data on the unit and approached by a research assistant for determination of interest in the study. Informed consent was obtained from both patient and caregiver together in a face-to-face meeting. The process of determining ability to consent for research followed guidelines for persons with cognitive

impairment (Alzheimer's Association, 2005) and was approved by the local ethics board. The recruitment period spanned one year (2008-9).

Data collection instruments

1) The Cognitive Performance Test (CPT)(Burns, 1992): The CPT is designed to observe the performance of persons with dementia on specific ADL and instrumental activities of daily living (IADL) tasks that require information processing skills. The individual performs up to seven activities (Medbox, Dress, Shop, Toast, Phone, Wash, and Travel)(approximately 45 minutes). Scoring is based on Allen Cognitive Level Theory for the level of cueing and demonstration required (1 to 6). Scores on the tasks are averaged to determine the total score. The test is designed for administration by an OT, and does not require specialized training.

2) The Assessment of Motor and Process Skills (AMPS)(Fisher, 2006): The purpose of the AMPS is to observe the performance of daily living tasks of the client's choice (negotiated after approximately 15 minute interview), for a wide range of ages and diagnostic groups. The client is observed performing two tasks (approximately one hour) and performance is scored on 15 motor (e.g. lifts, walks) and 20 process (e.g. initiates, sequences) skills to obtain two overall scores: motor and process score. Administration requires OT qualifications, AMPS training and certification, and was completed by the principal investigator.

3) Functional Independence Measure (FIM™)(Uniform Data System for Medical Rehabilitation, 1994): The patient's level of independence in activities of daily living and burden of care was observed over the course of caregiving in hospital. Scores

are assigned from 1 to 7 based on the level of independence. FIM scores are completed as a part of routine measurement on the rehabilitation unit and were obtained from the patient's medical chart.

4) Comorbidities: Cumulative Illness Rating Scale- Geriatric (CIRS-G) (Miller & Towers, 1991). The rating is determined by calculating a severity index. The index is the severity rating (based on mortality data) for each diagnosis provided in the manual, divided by the total number of categories endorsed. This measure has demonstrated good evidence for reliability and validity in comparison with other measures (Hall, 2006) and has data to support criterion validity for mortality rates (Conwell, Forbes, Cox & Caine 1993; Waldman & Potter 1992).

5) Incident of harm: An incident of harm was defined as (a) physical injury to self (including exacerbation of a chronic illness) or other, property loss, or property damage; (b) occurrence of the incident due to self-neglect or behaviours related to disorientation. This definition employed two of three criteria defined in a previous study (Tierney et al., 2004), but did not require the incident to have precipitated a medical visit. The incident of harm outcome was measured monthly with a caregiver questionnaire adapted from Tierney et al. 2004 (Appendix 1).

Data collection procedures

The CPT was administered to consenting participants by the attending OT and an independent assessor blind to the CPT score administered the AMPS. The independent variables age, sex, education, comorbidities and FIM score were collected at baseline from chart review by the PI or designate. Incident of harm

outcome data were gathered with consent of participants through a monthly telephone interview with the caregiver (Incident of Harm Caregiver Interview, Appendix 1). If the caregiver identified that an incident of harm involved hospitalization or a physician visit, confirmation was requested from the family physician and/or hospital medical records with the consent of participants (Appendix 3).

Analyses

The demographic characteristics of the sample are described. Demographic characteristics of persons who had an incident of harm at some point during the 6-month follow-up were compared with those who did not. The statistical significance of the difference in characteristics between those who did and did not have an incident of harm was analyzed using independent samples t-tests, chi squared analyses and effect size for chi square (w) or for t-test (Cohen's d using an average variance).

Objectives 1) and 2): Validity of CPT and AMPS: The ability of CPT score, AMPS motor and, AMPS process score (independent variables) to predict incident of harm (dependent variable) was analyzed in two ways. The CPT, AMPS-motor and AMPS-process scores were correlated with incident of harm outcome (Pearson's r). Second, the ability of the cut-scores to predict incident of harm was determined by calculating sensitivity (number scored as impaired/ number with incident of harm) and specificity (number scored unimpaired/number without incident of harm).

Objective 3): Contribution of independent variables: Cox proportional hazards

regression analyses were used to examine whether independent variables (age, sex, education, co-morbidities, ADL (FIM) score, caregiving hours) increased prediction of time to incident of harm over and above prediction by AMPS-P, AMPS-M or CPT. Time to incident of harm was measured in months from the date of discharge to the first month an incident was reported. Persons who did not have an incident or who were lost to follow-up were treated as censored observations. Analyses were performed separately for AMPS and CPT because of the clinical need to understand the performance of each measure individually. Independent variables were not able to be chosen based on theoretical grounds. Therefore it was important to examine the contribution of all variables to prediction of outcome using a single step. This minimized the chance of generating overly optimistic results by amplifying chance variation through step-wise model building (Babyak, 2004; Hosmer, Lemeshow & May, 2008). The number of variables combined in a single Cox regression equation was restricted to one variable for every 10 events (Hosmer, Lemeshow & May, 2008).

To test the influence of other variables on the predictive ability of the functional measures, the estimate of likelihood ratio for the equation with AMPS-P, AMPS-M or CPT alone was compared with estimates of likelihood for all possible two-variable equations combining the AMPS-P, AMPS-M or CPT with an independent variable. Variables were tested for covariance with AMPS-P, AMPS-M and CPT and excluded if they were significantly correlated ($p < 0.05$). The difference in likelihood ratio ($-2 \log$ likelihood) was analyzed by examining the statistical significance of the

Chi-squared statistic. To test whether the data met the proportional hazards assumption, Schoenfeld residuals and log minus log residuals were examined. Split sample reliability could not be calculated due to the number of events ($n=22$). The analyses were completed using PASW 17 statistical software.

Results

The sample at baseline ($n=47$) was composed of 55% females with a mean age of 83.5 years including 14 participants over the age of 85. A total of 45 persons contributed follow-up data; two participants dropped out after baseline assessment by declining to answer or return follow-up calls. Of the 45 persons, 39 completed six months of follow-up, two were deceased at follow-up, two caregivers dropped out while consenting for their information to remain in the study, and two became ineligible due to admission to long-term care facilities.

Persons who experienced an incident of harm at any time during the six-month follow-up differed from those who did not in two characteristics: living alone at discharge ($p<0.05$) and lower AMPS-process score ($p<0.01$) (Table 1).

Differences in CPT score approached significance ($p=0.05$) as did AMPS-M scores and paid caregiving hours. The groups did not differ with statistical significance across other characteristics such as age, sex, education, FIM score, co-morbidities, or caregiving hours. Effect size estimates demonstrate that the effect of AMPS- process was large and of CPT was moderate. There were also moderate effects of age, paid caregiving hours and AMPS-motor score. The types of incidents were classified by

asking the caregiver to estimate the most likely reason for the incident of harm.

Responses were grouped into categories according to the method of Tierney (2004).

All incidents in all months were coded (n= 35) The highest proportion of incident was of the type “failure to use mobility devices correctly” (n=11, 31.4%), followed by “failure to use medication correctly” (n=7, 20.0%), “failure to perform ADL” (n=6, 17.1%), “failure to report medical condition” (n=6, 17.1%), and “unknown” type (n=5, 14.3%).

Table 1: Comparison of participants who did and did not experience harm

Characteristic	Incident of harm (n= 22)		No incident of harm (n=23)		Analysis			
	mean	SD	mean	SD	parameter	df	p	Effect size†
Age (years)	85.26	7.85	81.41	7.52	t= -1.68	43	0.10	d=0.50
Sex	male=8 female=15		male=13 female=9		$\chi^2=2.67$	1	0.10	w=0.24
Education (years)	10.43	3.19	9.86	3.19	t= -0.61	43	0.54	d=0.18
Lives alone at discharge	yes=11 no=12		yes=4 no=18		$\chi^2=4.45$	1	0.04*	w=0.33
Functional Independence Measure	100.91	11.56	106.09	11.54	t= 1.50	43	0.14	d=-0.45
Co-morbidity index (CIRS-G)	2.05	0.23	2.00	0.30	t= -0.65	43	0.52	d=0.19
Paid caregiving hours	8.10	8.56	3.98	6.57	t=-1.81	43	0.08	d=0.54
Unpaid caregiving hours	19.57	17.04	15.96	10.54	t=-0.85	43	0.40	d=0.25
Total caregiving hours	27.67	18.04	19.89	13.00	t= -1.66	43	0.10	d=0.49
AMPS-process	0.50	0.46	0.86	0.39	t= 2.83	43	<0.01**	d=-0.84
AMPS-motor	0.83	0.53	1.12	0.52	t= 1.85	43	0.07	d=-0.55
CPT	4.44	0.42	4.65	0.28	t= 2.01	43	0.05	d=-0.59
SMMSE	23.17	4.72	24.27	3.41	t= 0.89	43	0.38	d=-0.27

† Effect size small d=0.2, moderate d=0.5, large d=0.8; small w=0.1, medium w=0.3

* significant at p<0.05 level

** significant at p<0.01 level

SMMSE: Standardized Mini Mental Status Exam

AMPS- motor: Assessment of Motor and Process Skills- motor scale

AMPS- process: Assessment of Motor and Process Skills- process scale

CPT: Cognitive Performance Test

Validity of AMPS and CPT

The AMPS-process scale showed a statistically significant correlation with time to first incident of harm (Pearson's $r = 0.37$, $p = 0.02$) ($n = 44$). Non-significant correlations were found for the AMPS-motor scale ($r = 0.29$, $p = 0.05$) and CPT ($r = 0.23$, $p = 0.13$) although the AMPS-motor scale approached significance.

The accuracy of the cut-points for identifying persons with an incident of harm was analyzed for sensitivity and specificity (Table 3). The AMPS cut-points are pre-established and described in the test manual. The CPT test manual does not specify a cut-point, but for purposes of analysis of measurement performance, a cut point of 4.2 was designated based on evidence cited in the test manual (Burns, 1992) that this cut-point (4.2 or lower) predicted nursing home admission.

When examining specificity, the most accurate measure was the CPT (Table 3). The CPT was accurate when identifying persons who did not have an incident of harm, with a specificity of 0.91. In other words, for persons scoring above 4.2, the CPT identified that the person did not have an incident of harm with 91% accuracy. Sensitivity (ability to accurately predict persons with an incident of harm) was low for all measures, but was greatest for the AMPS-P scores with 67% of those with an incident having AMPS-P scores below the cut-point.

Table 2: Sensitivity and specificity of functional measures for incident of harm

Measure	Sensitivity (incident of harm) (n=23)	Specificity (no incident of harm) (n=21)
CPT (cut-point 4.2)	0.44	0.91
AMPS-motor	0.55	0.70
AMPS-process	0.67	0.68

Note: Sensitivity= #impaired with incident of harm/ # with incident of harm
 Specificity= # not impaired without incident of harm/ # without incident of harm

Contribution of independent variables

The data adhered to the assumptions of Cox regression analysis as plots of Schoenfeld residuals against time to incident of harm showed random values (r squared less than 0.05).

To examine if any single variable increased the ability to predict time to incident of harm over and above the AMPS-P, AMPS-M or CPT, Cox regression analysis was performed. Single variable analysis estimated the contribution of each variable to predict incident of harm. Table 3 shows an estimate of the direction and size of the effect (hazard ratio), and uncertainty of the effect (confidence interval). The AMPS-process scale was the best single predictor of time to incident of harm (p=0.01). The CPT was the next best predictor and approached statistical significance.

Hazard ratios, shown in Table 3, describe the change in time to incident of harm for every one-point change in a variable. Hazard ratios below 1.0 indicate decrease in the outcome with increase in the independent variable, and above 1.0

indicate increase in the outcome. The direction of change in time to incident of harm was as follows: an increase in AMPS-process score corresponded to a decrease in incident of harm hazard ratio. The direction of the CPT hazard ratio was the same: an increase in CPT score corresponded to a decrease in incident of harm hazard ratio. The value of the hazard ratio cannot be directly compared between the measures because each measure has a scale with a different range.

Several other variables approached statistical significance in predicting of time to incident of harm. After AMPS-P and CPT, incidents of harm were best predicted by higher AMPS- Motor skills score (p=0.07), living alone (p=0.07), higher total caregiving hours (p=0.08), increased age (p=0.11), and being female (p=0.12). The estimates of hazard ratios showed wide confidence intervals.

Table 3 Cox regression analysis predicting time to harm: single variable analysis (n=45)

Variable	Variable p-value	Hazard ratio (95% CI)
AMPS-process	0.01*	0.31 (0.13-0.77)
AMPS-motor	0.07	0.48 (0.21-1.07)
CPT	0.06	0.31 (0.09-1.05)
Age	0.11	1.05 (0.99-1.11)
Sex (female vs. male)	0.12	2.00 (0.85-4.74)
Education	0.51	1.05 (0.91-1.21)
Lives alone	0.07	0.46 (0.20-1.06)
FIM	0.15	0.98 (0.94-1.01)
Co-morbidities	0.48	1.74 (0.38-8.10)
SMMSE	0.43	0.96 (0.87-1.06)
Paid caregiving hours	0.10	1.04 (0.99-1.08)
Unpaid caregiving hours	0.30	1.01 (0.98-1.05)
Total caregiving hours	0.08	1.02 (1.00-1.05)

* significant at p<0.05

To determine whether adding any independent variable to AMPS-P, AMPS-M

or CPT increased the prediction of time to incident of harm, two-variable Cox regression equations were examined. Estimates of overall predictive ability for all two-variable equations were analyzed in comparison to the estimate with AMPS-P, AMPS-M or CPT only. Table 4 shows that, for this sample, AMPS-P and living alone increased the estimate of hazard for time to incident of harm compared to the estimate with AMPS-P alone ($\chi^2 = 4.32, p=0.04$). Other variables that approached significance were sex ($\chi^2 = 3.64, p=0.06$) and age ($\chi^2 = 2.22, p=0.14$). Estimates of hazard ratios showed wide confidence intervals.

Living alone contributed significantly to overall predictive ability of AMPS-M (Chi square change = 5.22, $p=0.02$). In the two-variable model, the hazard ratio for AMPS-M was 0.41 (CI: 0.19-0.88) and for living alone was 2.74 (CI: 1.18-6.37). The bivariate model was a significant predictor of incident of harm outcome (Chi square= 9.7, $p<0.01$).

Table 4: Cox regression model: contribution of independent variables in predicting time to harm with AMPS-P (n=45)

Variables in regression model	Hazard ratio (95% CI)	Chi square compared to AMPS-P only (Df=1)	p-value
AMPS-P	0.34 (0.14- 0.82)	2.22	0.14
Age	1.04 (0.99- 1.11)		
AMPS-P	0.27 (0.11- 0.70)	3.64	0.06
Sex	2.29 (0.96- 5.48)		
AMPS-P	0.32 (0.12- 0.81)	0.01	0.96
Years education	1.00 (0.87-1.17)		
AMPS-P	0.27 (0.11- 0.70)	4.32	0.04*
Lives alone	2.47 (1.07- 5.71)		
AMPS-P	0.32 (0.12- 0.82)	0.02	0.89
Co-morbidities (CIRS-G)	1.12 (0.21- 5.87)		
AMPS-P	0.34 (0.13- 0.88)	0.48	0.49
FIM	0.99 (0.95- 1.02)		
AMPS-P	0.34 (0.13- 0.85)	1.78	0.18
Total caregiving hours	1.02 (0.99- 1.05)		
AMPS-P	0.35 (0.13- 0.90)	0.61	0.44
Paid caregiving hours	1.02 (0.97- 1.07)		
AMPS-P	0.31 (0.13- 0.77)	0.99	0.32
Unpaid caregiving hours	1.02 (0.99- 1.05)		
AMPS-P	0.36 (0.13- 0.97)	0.87	0.35
AMPS-Motor	0.66 (0.28- 1.58)		
AMPS-P	0.28 (0.10- 0.78)	0.25	0.62
SMMSE	1.03 (0.92- 1.15)		

* significant at p<0.05

Table 5: Cox regression model: contribution of independent variables in predicting time to harm with AMPS- M (n=45)

Variables in regression model	Hazard ratio (95% CI)	Chi square compared to AMPS-P only (Df=1)	p-value
AMPS-M	0.47 (0.20- 1.06)	2.89	0.09
Age	1.05 (0.99-1.11)		
AMPS-M	0.51 (0.22-1.17)	1.92	0.17
Sex	1.82 (0.77-4.3)		
AMPS-M	0.49 (0.21- 1.12)	0.05	0.82
Years education	1.02 (0.88-1.18)		
AMPS-M	0.41 (0.19-0.88)	5.22	0.02*
Lives alone	2.74 (1.18-6.37)		
AMPS-M	0.49 (0.22-1.11)	0.18	0.68
Co-morbidities (CIRS-G)	1.40 (0.29-6.71)		
AMPS-M	0.52 (0.17-1.55)	0.05	0.83
FIM	0.10 (0.95-1.04)		
AMPS-M	0.55 (0.23-1.32)	1.34	0.25
Total caregiving hours	1.02 (0.99-1.04)		
AMPS-M	0.55 (0.24-1.26)	0.95	0.33
Paid caregiving hours	1.02 (0.98-1.07)		
AMPS-M	0.49 (0.21-1.14)	0.50	0.48
Unpaid caregiving hours	1.01 (0.98-1.04)		
AMPS-M	0.48 (0.21-1.09)	0.46	0.50
SMMSE	0.97 (0.88-1.06)		

Table 6: Cox regression model: contribution of independent variables in predicting time to harm with CPT (n=45)

Variables in regression model	Hazard ratio (95% CI)	Chi square compared to CPT only (Df=1)	p-value
CPT	0.22 (0.06- 0.77)	4.81	0.03*
Age	1.07 (1.00- 1.14)		
CPT	0.27 (0.07- 0.96)	3.23	0.07
Sex	2.17 (0.91- 5.20)		
CPT	0.31 (0.09- 1.04)	0.49	0.48
Years education	1.05 (0.91- 1.21)		
CPT	0.29 (0.09- 1.00)	3.54	0.06
Lives alone	2.24 (0.98- 5.11)		
CPT	0.32 (0.09- 1.11)	0.18	0.67
Co-morbidities (CIRS-G)	1.40 (0.30- 6.64)		
CPT	0.38 (0.10- 1.37)	0.63	0.43
FIM	0.99 (0.95- 1.02)		
CPT	0.37 (0.10- 1.30)	1.02	0.31
Paid caregiving hours	1.02 (0.98- 1.07)		
CPT	0.34 (0.09- 1.22)	0.17	0.68
Unpaid caregiving hours	1.01 (0.98- 1.04)		
CPT	0.32 (0.10-1.08)	3.11	0.08
AMPS-Motor	0.47 (0.20- 1.11)		
CPT	0.27 (0.06- 1.20)	0.11	0.74
SMMSE	1.02 (0.91- 1.15)		

For this sample, CPT and age statistically significantly increased the estimate of hazard for time to incident of harm compared to the estimate with CPT alone (Table 6). This bivariate model with CPT and age was a statistically significant predictor of incident of harm outcome (Chi square =8.4, $p < 0.05$). The next most predictive variables when combined with CPT were living alone ($\chi^2 = 3.54$, $p = 0.06$)

and sex ($\chi^2 = 3.23, p=0.07$). The confidence intervals around the hazard ratio estimates were wide.

All variables were examined for correlations in order to describe the results. A covariance matrix was generated (Table 7). A variable that is correlated to either AMPS-P or CPT will be less likely to explain the total variance in the outcome incidents of harm, thus is less likely to show a statistically significant chi square value. Variables correlated with AMPS-P were: FIM ($r= 0.33, p<0.05$), AMPS-M ($r=0.44, p<0.01$) and SMMSE ($r=0.46, p<0.01$). Variables correlated with CPT were: FIM ($r=0.32, p<0.05$), total caregiving hours ($r= -0.48, p<0.01$), and SMMSE ($r=0.47, p<0.01$). The variables living alone and age were not correlated with AMPS-P, AMPS-M or CPT.

Table 7 Covariance matrix: Pearson's r correlation of all variables

Variable Correlation	Age	Sex	Lives alone	Years education	FIM	CIRS Index	Total caregiving hours	Paid caregiving hours	Unpaid caregiving hours	SMMSE	AMPS- Motor	AMPS-Process	CPT
Age	1	.07	.30*	.18	.17	-.11	-.31*	-.08	-.30*	.01	.06	-.03	.12
Sex		1	.16	.09	-.02	.14	-.07	-.04	-.06	-.07	-.19	-.01	-.02
Lives alone			1	-.01	.28	-.32*	-.40**	.08	-.50**	.26	.20	.09	.15
# years education				1	-.24	-.12	.03	-.01	.03	.04	-.34*	-.22	-.06
FIM					1	-.09	-.58**	-.39**	-.42**	.19	.62**	.33*	.32*
CIRS index						1	.11	.26	-.02	-.38**	-.28	-.22	-.26
Total caregiving hours							1	.47**	.87**	-.14	-.30*	-.26	-.48**
Paid caregiving hours								1	-.02	-.05	-.27	-.27	-.25
unpaid caregiving hours									1	-.12	-.19	-.14	-.41**
SMMSE										1	.11	.46**	.47**
AMPS-Motor											1	.43**	.15
AMPS-Process												1	.53**
CPT													1

* significant at p<0.05 ** significant at p<0.01

Discussion

The overall purpose of the study was to examine the predictive validity of CPT and AMPS. The outcome against which they were tested was incidents of harm in the community. Validation of these measures for incident of harm outcome is important because data are needed to determine whether incidents of harm in the person's home environment can be predicted using measures administered in hospital. This allows accurate interpretation and ethical use of the measures.

The AMPS-Process scale was the best single predictor of time to incident of harm outcome. A low score on AMPS-Process was related to shorter time to incident of harm. With regard to the CPT, its strength was that the specificity (the chance of identifying a person who will not have harm) was higher than for any other measure. The two functional measures of cognition, CPT and AMPS-Process were more predictive of harm in this sample than other clinical measures such as activities of daily living (FIM), and co-morbidities (CIRS-G). They were also more predictive than a common cognitive screening tool (SMMSE). The relationship of the AMPS-Process to incident of harm outcome and the CPT specificity, combined with the lack of relationship found for other measures, provides preliminary argument for the validity of the AMPS-Process and CPT for prediction of harm after discharge.

When comparing motor and cognitive measures, AMPS-Motor scale showed less predictive ability for incident of harm outcome than the functional measures of cognition. Living alone contributed significantly to the ability of motor skills scores to predict incident of harm outcome, as it did with AMPS-Process. AMPS-Process and

AMPS-Motor scores are obtained from observation of the same assessment tasks as they are administered and scored simultaneously. Thus, AMPS-Motor scores would always accompany AMPS-Process scores. Since AMPS-Process had stronger predictive ability than AMPS-Motor, and this was not significantly improved by adding AMPS-Motor in the predictive model, the data in this study support the preference of the AMPS-Process score over the AMPS-Motor as an indicator of incidents of harm outcome.

The case for the validity of cognitive measures for predicting harm is supported when examining data from this study in light of data from previous studies. Data from several sources gave preliminary indication that measures of cognition may have predictive validity for incidents of harm. The AMPS-Process scale was correlated with a home safety checklist for persons with psychiatric conditions (McNulty & Fisher, 2001). The AMPS-P scale had stronger correlation with the safety checklist than the AMPS-Motor scale (n=20). The limitation of this previous study was the uncertainty about the correlation of the safety checklist with actual incidents of harm. The contribution of this current study is it provides stronger evidence for the ability of the AMPS-Process to predict actual incidents of harm.

Other neuropsychological measures also have been correlated with incidents of harm (Tierney, 2007). The limitation of the previous data was that the sample was drawn only from persons who lived alone. As well, the definition of an “incident of harm” was more stringent in that only an incident that required medical

attention was noted, which limited the types of incidents against which the measures could be validated (Tierney, 2007). The evidence from this current study adds to the literature by demonstrating cognitive measures can be predictive of incidents resulting in minor injury not seen by the physician but known to the caregiver, and predictive for a population that includes both persons who live alone or with others.

Furthermore, the following accumulated evidence associates cognition with increased morbidity. The severity of dementia, as measured by the Clinical Dementia Rating or MMSE, was a significant risk factor for nursing home admissions (Rozzini et al., 2001, Tun et al., 2007), repeat emergency department visits (McKusker et al., 1997) and hospitalization independent of age, gender, education, and co-morbid medical conditions (Albert et al., 1999). Risk of accidental falls is increased with decreasing cognitive status as an independent predictor (Anstey, Von sadden & Luszcz, 2006; Gleason, Gagnon, Fischer & Mahoney, 2009; Hausdorff & Yogev, 2006; Liu-Ambrose, Pang & Eng, 2007; Rapport, Hanks, Millis & Deshpande, 1998) or in combination with other factors such as balance, strength and nutritional status (Chen et al., 2005; Stolee, Poss, Cook, Byrne & Hirdes, 2009). It is suggested that the reason for the association between cognitive status and falls is that persons with cognitive impairment behave in ways that precipitate falls, or that at the least are not effective in preventing falls (Hausdorff & Yogev, 2006). This evidence, taken together with data from this study suggests that understanding the cognitive status of an older adult informs an understanding of risk for incidents of harm. This study

provides initial evidence for specific measures that directly relate to incidents of harm at home.

The clinical interpretation of the results of this study is that scores on the measures can be used to inform patients, families and the team about risk of incidents of harm after discharge. The evidence suggests that a low score on the AMPS-Process in hospital was associated in this sample with incidents of harm at home within six months after discharge. A low score on CPT has a weak association with such incidents, but a CPT score above the cut-score of 4.2 may indicate low risk of harm outcome. Caution regarding test interpretation arises when examining the data. The cut-point for the AMPS-Process was poor at accurately predicting harm outcomes. A score below the cut-point for AMPS-P or CPT predicted the likelihood of having an incident of harm with approximately 70% accuracy (sensitivity). This limited precision is important to note when interpreting a score. The cut-point for the CPT was more accurate than the AMPS-Process at identifying persons who did not have harm (91% specificity). A precise statement of the incremental change in risk for each score (hazard ratio) is not supported by the data because of the wide confidence intervals noted in this study. The data in this study demonstrated weak to moderate correlations (Pearson's $r < 0.5$) between AMPS-P, CPT, AMPS-M and incident of harm outcome, indicating that multiple factors may contribute to incidents of harm.

The data support the clinical use of the measures to detect the possibility of harm in this sample but indicate that several other independent variables also

contributed to the prediction of harm and should be included in clinical considerations of risk for harm outcome. The t-test results (Table 1) show that living alone was attributed to increased probability of having at least one incident of harm. As well, although statistically non-significant, there was moderate effect of age, paid and total caregiving hours and AMPS-motor scores (Table 1). The results of the Cox regression analysis showed that living alone significantly increased the predictive ability of the AMPS-P for time to incident of harm. Age, sex and living alone had the greatest effect on the predictive ability of either AMPS-P or CPT. Because the predictive ability of the AMPS-P and CPT was influenced by these variables, this means that an individual's score cannot be interpreted in isolation from other variables in that a low score on the measure combined with living alone, greater age and being female in this sample was associated with increased risk for incident of harm outcome during the six months after discharge.

When comparing the clinical feasibility of the AMPS-Process and CPT, consideration must be given to the relative precision to predict harm as well as the training, administration time and purchase cost. The AMPS requires a large investment since it has a required weeklong training course and software purchase. The CPT requires purchase of the test kit and self-training with the test manual. It is unlikely that a clinician would administer both AMPS and CPT because of time constraints. The advantage of the AMPS was statistically stronger prediction for the incident of harm outcome. The CPT was useful because of the high specificity, thus the ability to detect persons who did not have harm, but in this study, was a weaker

predictor of harm overall than AMPS. It is also interesting to note that the cognitive screening measure SMMSE showed comparatively lower correlation with incident of harm outcome than either the CPT or AMPS. It can be hypothesized that the cognitive screen was not sufficiently sensitive to differences in cognition that may lead to accidental injury. The SMMSE is poor at identifying mild cognitive impairment (Patterson et al. 2001) and the narrow range of scores in this study may have limited our ability to find association with the outcome. In this study, the CPT and AMPS-Process were superior to the SMMSE. Therefore including the results of one of these measures rather than the SMMSE alone in discharge planning may be advisable to best care of the patient and family.

Additionally, this study provides data supporting standardized measures. Although attractive to clinicians because of low cost, and adaptability, non-standardized functional observations such as cooking assessments, because of their lack of standardization, can have high variability in administration or interpretation. Moreover, the degree of variability is not known. Thus, this study contributes to evidence-based practice by providing preliminary data where it is lacking for non-standardized assessments. This study demonstrated that the inclusion of a standardized functional measure has value to determine the degree of increased risk, and provide a description of contributing factors age, sex and living situation, as well as the precision with which harm outcomes could be predicted. These data can be used in the best practice of occupational therapy to support ethical discharge planning and decision-making for patients and families.

Limitations and directions for future research:

Study limitations were related to sources of measurement error and enrollment of participants. Effort was made to reduce measurement bias. Measurement with the CPT, FIM and SMMSE may have been associated with greater variability than AMPS because, for protocol feasibility, they were part of usual care and therefore administered by different raters. This could lower the inter-rater reliability, and thereby lower the correlations observed for the CPT compared to the AMPS, which was administered by a single rater. That said, raters were trained in an effort to maximize inter-rater reliability. Further, reliability and validity of the incident of harm outcome was determined for this sample as discussed in further detail elsewhere (Douglas, Letts, Richardson & Eva, manuscript, thesis Chapter 3). When checked against medical charts, the caregiver reports were accurate as previously described (Douglas, Letts, Richardson & Eva, manuscript, thesis Chapter 3).

The number of patient and caregiver pairs that were enrolled limited the analyses and the number of incidents of harm limited the Cox regression analysis. Margins of error associated with the analyses are provided and a larger sample size would have reduced the error. The range of scores was restricted in this sample and greater heterogeneity would reduce the possibility of Type II error (finding of no correlation when one exists) for variables such as SMMSE, ADL burden of care of motor score. Challenges related to sample size included the time available for the study, the need for a six-month follow up and ethical considerations for persons

with dementia necessitated obtaining consent from both patient and a caregiver. The advantage of recruitment over a full year was that seasonal bias for incidents of harm (e.g. slips on ice) was less likely to occur. In the future, researchers must consider recruitment from more than one hospital unit, perceived caregiver burden and longer study time in increasing the power and generalizability of the results. Further evidence based on a larger sample size is needed to corroborate the validity argument and determine whether more precise estimates of the risk associated with obtaining a score above or below a certain cut-point for each measure can be obtained.

Conclusion

Functional measures of cognition were validated for predicting incident of harm outcome. The AMPS-P and CPT were found in this study to aid in identifying persons at risk for harm. Moreover, it was found that measures administered in the hospital predicted an outcome after discharge. Other variables including living alone, age and being female were also found to influence the ability of the measures to predict incident of harm outcome. When identifying a person at risk, interpretation must be accompanied by a caveat that the precision of cut scores is limited. Individual in-patient scores, when combined with knowledge of living situation, sex and age, give an indication of risk for incidents of harm.

This study has both research and clinical applications. The scholarly contribution is the validation of in-hospital measures against a real-world outcome that quantifies harm in the community. It contributes clinically by providing the

first evidence to support the interpretation of the scores on both measures for discharge planning, but also by supporting the use of these measures for predicting incidents of harm over other measures that were less predictive such as co-morbidities and ADL (FIM). The predictive validity of these functional measures of cognition provides important evidence to further develop measures for assisting with planning for older adults' safety in the community.

References

- Albert, S. M., Costa, R., Merchant, C., Small, S., Jenders, R. A., & Stern, Y. (1999). Hospitalization and Alzheimer's disease: Results from a community-based study. *Journals of Gerontology Series A-Biological Sciences & Medical Sciences*, *54*, M267-71.
- Alzheimer's Association. (2004). Research consent for cognitively impaired adults: Recommendations for institutional review boards and investigators. *Alzheimer Disease and Associated Disorders*, *18*, 171-175.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for Educational and Psychological Testing*. Washington, DC: American Psychological Association.
- Anstey, K. J., von Sanden, C., & Luszcz, M. A. (2006). An 8-year prospective study of the relationship between cognitive performance and falling in very old adults. *Journal of the American Geriatrics Society*, *54*, 1169-1176.
- Babiyak, M. A. (2004). What you see may not be what you get: A brief, nontechnical introduction to over-fitting in regression-type models *Psychosomatic Medicine*, *66*, 411- 421.
- Banaszak-Holl, J., Fendrick, A. M., Foster, N. L., Herzog, A. R., Kabeto, M. U., Kent, D. M., et al. (2004). Predicting nursing home admission: Estimates from a 7-year follow-up of a nationally representative sample of older Americans. *Alzheimer Disease & Associated Disorders*, *18*, 83-89.
- Bar-Yosef, C., Weinblatt, N., & Katz, N. (1999). Reliability and validity of the Cognitive Performance Test (CPT) in an elderly population in Israel. *Physical and Occupational Therapy in Geriatrics*, *17*, 65-79.
- Burns, T. (1992). The Cognitive Performance Test: An approach to cognitive level assessment in Alzheimer disease. In C. K. Allen, C. A. Earhart, & T. Blue (Eds.), *Occupational Therapy Treatment Goals for the Physically and Cognitively Disabled* (pp. 46-84). Rockville, MD: The American Occupational Therapy Association, Inc.
- Chen, J. S., March, L. M., Schwarz, J., Zochling, J., Makaroff, J., Sitoh, Y. Y., et al. (2005). A multivariate regression model predicted falls in residents living in intermediate hostel care. *Journal of Clinical Epidemiology*, *58*, 503-508.

- Conwell Y., Forbes N.T., Cox, C., Caine, E.D. (1993). Validation of a measure of physical illness burden at autopsy: the Cumulative Illness Rating Scale. *Journal of the American Geriatric Society*, 41, 38–41.
- Cooper, S. (1981). Accidents and older adults. *Geriatric nursing*, 2, 287-290.
- Corcoran, M. (2001). Dementia. In B.R. Bonder & M.B. Wagner (Eds.) *Functional Performance in Older Adults* (2nd ed.) pp. 287-304.
- Douglas, A., Letts, L., & Liu, L. (2008). Review of cognitive assessments for older adults. *Physical and Occupational Therapy in Geriatrics*, 26(4), 13-43.
- Douglas, A., Liu, L., Warren, S., & Hopper, T. (2007). Cognitive assessments for older adults: Which ones are used by Canadian occupational therapists and why. *Canadian Journal of Occupational Therapy*, 74, 370-381.
- Dyer, C. B., Pickens, S., & Burnett, J. (2007). Vulnerable elders: When it is no longer safe to live alone. *Journal of the American Medical Association*, 298, 1448.
- Fisher, A. G. (2006). *Assessment of Motor and Process Skills. Vol. 2: User Manual* (6th ed.) Fort Collins, CO: Three Star Press.
- Gleason, C. E., Gangnon, R. E., Fischer, B. L., & Mahoney, J. E. (2009). Increased risk for falling associated with subtle cognitive impairment: Secondary analysis of a randomized clinical trial. *Dementia & Geriatric Cognitive Disorders*, 27, 557-563.
- Hall, S.F. (2006) A user's guide to selecting a comorbidity index for clinical research. *Journal of Clinical Epidemiology* 59, 849–855.
- Hausdorff, J. M., & Yogev, G. (2006). Cognitive function may be important for fall injury prevention trials. *Journal of the American Geriatrics Society*, 54, 865.
- Hosmer, D.W., Lemeshow, S., & May, S. (2008). *Applied survival analysis: Regression modeling of time to event data*. Hoboken, NJ: Wiley Series in Probability and Statistics.
- Jonsson, L., Eriksson, M., Kilander, L., Soininen, H., Hallikainen, M., Waldemar, G., et al. (2006). Determinants of costs of care for patients with Alzheimer's disease. *International Journal of Geriatric Psychiatry*, 21, 449-459.
- Kane, M. T. 2006. Validation. In R. Brennan (Ed.) *Educational Measurement, 4th edition*. Wesport, CT: American Council on Education and Praeger, 17 - 64.

- Lilley, J. M., Arie, T., & Chilvers, C. E. (1995). Accidents involving older people: A review of the literature. *Age & Ageing, 24*, 346-365.
- Lindén, A., Boschian, K., Eker, C., Schalén, W., & Nordström C.H. (2005). Assessment of Motor and Process Skills reflects brain-injured patients' ability to resume independent living better than neuropsychological tests. *Acta Neurologica Scandinavica, 111*, 48-53.
- Liu-Ambrose, T., Pang, M. Y., & Eng, J. J. (2007). Executive function is independently associated with performances of balance and mobility in community-dwelling older adults after mild stroke: Implications for falls prevention. *Cerebrovascular Diseases, 23*, 203-210.
- McNulty, M. C., & Fisher, A. G. (2001). Validity of using the Assessment of Motor and Process Skills to estimate overall home safety in persons with psychiatric conditions. *American Journal of Occupational Therapy, 55*, 649-655.
- McCusker, J., Bellavance, F., Cardin, S., Belzile, E., & Verdon, J. (2000). Prediction of hospital utilization among elderly patients during the 6 months after an emergency department visit. *Annals of Emergency Medicine, 36*, 438-445.
- Messick, S. (1995). Validity of psychological assessment. *American Psychologist, 50*, 741-749.
- Miller, M. D., & Towers A. (1991). *Cumulative Illness Rating Scale for Geriatrics Manual*. University of Pittsburgh: Pittsburgh, PA.
- Mokkink, L. B., Terwee, C. B., Knol, D. L., Stratford, P. W., Alonso, J., Patrick, D. L., et al. (2006). Protocol of the COSMIN study: CONsensus-based standards for the selection of health measurement INSTRUMENTs. *BMC Medical Research Methodology, 6*, 2.
- Norman, G. R. & Streiner, D. L. (2000). *Biostatistics: The bare essentials* (2nd ed.). Hamilton ON: B.C. Decker Inc.
- Nasreddine, Z.S., Phillips, N.A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J.L., Chertkow, H. (2005). The Montreal Cognitive Assessment (MoCA): A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society, 695-699*.
- Patterson, C., Gauthier, S., Bergman, H., Cohen, C., Feightner, J. W., Feldman, H., et al. (2001). The recognition, assessment and management of dementing

disorders: Conclusions from the Canadian consensus conference on dementia. *Canadian Journal of Neurological Sciences*, 28(Suppl 1), S3-16.

- Rapport, L. J., Hanks, R. A., Millis, S. R., & Deshpande, S. A. (1998). Executive functioning and predictors of falls in the rehabilitation setting. *Archives of Physical Medicine & Rehabilitation*, 79, 629-633.
- Smith, G. E., Kokmen, E., & O'Brien, P. C. (2000). Risk factors for nursing home placement in a population-based dementia cohort. *Journal of the American Geriatrics Society*, 48, 519-525.
- Streiner, D. L., & Norman, G. R. (2003). *Health measurement scales: A practical guide to their development and use* (3rd ed.). New York: Oxford University Press.
- Tierney, M. C., Charles, J., Naglie, G., Jaglal, S., Kiss, A., & Fisher, R. H. (2004). Risk factors for harm in cognitively impaired seniors who live alone: A prospective study. *Journal of the American Geriatrics Society*, 52, 1435-1441.
- Tierney, M. C., Snow, W. G., Charles, J., Moineddin, R., & Kiss, A. (2007). Neuropsychological predictors of self-neglect in cognitively impaired older people who live alone. *American Journal of Geriatric Psychiatry*, 15, 140-148.
- Tullis, A., & Nichol, M. (1999). A systematic review of the evidence for the value of functional assessment of older people with dementia. *British Journal of Occupational Therapy*, 62, 554-63.
- Tun, S. M., Murman, D. L., Long, H. L., Colenda, C. C., & von Eye, A. (2007). Predictive validity of neuropsychiatric subgroups on nursing home placement and survival in patients with Alzheimer disease. *American Journal of Geriatric Psychiatry*, 15, 314-327.
- Rozzini, L., Cornali, C., Chilovi, B. V., Ghianda, D., Padovani, A., & Trabucchi, M. (2006). Predictors of institutionalization in demented patients discharged from a rehabilitation unit. *Journal of the American Medical Directors Association*, 7, 345-349.
- Stolee, P., Poss, J., Cook, R. J., Byrne, K., & Hirdes, J. P. (2009). Risk factors for hip fracture in older home care clients. *Journals of Gerontology Series A-Biological Sciences & Medical Sciences*, 64, 403-410.
- Vassallo, M., Mallela, S. K., Williams, A., Kwan, J., Allen, S., & Sharma, J. C. (2009). Fall risk factors in elderly patients with cognitive impairment on rehabilitation wards. *Geriatrics & Gerontology International*, 9(1), 41-46.

Waldman E. & Potter, J.F. (1992). A prospective evaluation of the cumulative illness rating scale. *Aging*, 4, 171-178.

Appendix 1:
Incident of Harm Caregiver Interview
(reference: Tierney et al, 2004)

Follow up interview checklist: harm outcome

1. Name of patient participant

2. Outcome: Incidents of harm:

Can you tell me if any of the following have occurred with your __ (eg parent, spouse) in the past month?

__(a) any injury to self or another person (including exacerbation of a chronic illness), property loss, or property damage

__(b) occurrence of the incident due to self-neglect or behaviours related to disorientation (unintentional)

3. In your opinion, what might have been the **most important reason for the occurrence? (Type of incident)**

ADL:

Not completing ADL (eating, hygiene, dressing)

Not using medications correctly:

Not using mobility devices correctly:

IADL:

Not using kitchen or electrical devices correctly:

Not detecting fraud:

Not reporting a medical condition (e.g. medical condition exacerbated because did not report to home care, caregiver, or medical team):

Environment:

Not recognizing a familiar environment:

Unsafe home set-up:

Other:

4. If no incident of harm is reported, ask “So, is everything going OK?”

CHAPTER 5

Discussion

There is a lack of evidence validating measures used in occupational therapy with post-discharge outcomes. Moreover, there is a pressing need in health care to identify the most accurate measures to assist with discharge planning for older adults. Older adults, their families and the health care team require evidence to help support decisions about possible risks upon discharge and provision of supports for independent living. The overall purpose of the thesis was to evaluate the performance of the Assessment of Motor and Process Skills (AMPS) and Cognitive Performance Test (CPT) for pre-discharge assessment of hospital inpatients to aid in discharge planning. Two occupational therapy measures were studied. Both measures used a “top-down” approach by observing daily tasks to infer ability at a performance component level. A prospective study design was implemented to examine the validity of functional measures for predicting harm outcome. In the first manuscript, baseline data were examined to describe the CPT for internal consistency reliability and the way the measure identified impairment compared to other in-hospital measures. The second described the reliability and validity of a measure of harm outcome after hospital discharge. In the final paper, the measures were examined for their ability to predict the harm outcome.

Overall, the studies offer data to increase precision of clinical interpretation of the measures when used for hospital discharge planning for older adults. They

provide the first data to link in-hospital functional assessment with a harm outcome after discharge. As such, they add to the evidence for clinical interpretation when patients, families and health care teams are making decisions about services and being discharged home. This discussion will address the contributions of each manuscript individually, then discuss overall implications with regard to application of measurement theory and clinical implication of the validity evidence.

Contributions of each manuscript

CPT study

In the CPT study, validity assumptions were tested. There was a need for further validation of the CPT because, compared with the other functional measure (the AMPS), the CPT had only an adequate level of evidence describing various aspects of validity. The CPT also has potential to become widely used because of its lower purchase cost and minimal requirement for additional training, thus making the need more urgent amongst clinicians for data to enable accurate test score interpretation.

In this study, CPT scores were not significantly correlated with measures of motor skill and co-morbidity. They were correlated with measures of cognition. CPT scores were not influenced by age, sex or education. Together, this evidence supports the validity of the CPT as a measure of the construct of cognition.

Examination of the precision with which the CPT designated impairment (decision rule) found that it was similar to AMPS-P and SMMSE (significant phi

coefficients) but the degree of agreement was weak. In the absence of a “gold standard” criterion, this means that whether a person is identified as having impairment may depend on characteristics of the measure chosen by the therapist, but also on error associated with measurement. The implications of this study are that test results cannot be interpreted to be an absolute indication of a person’s level of independence in daily living. Thus, an important contribution of this study is to demonstrate what is *not* recommended for test interpretation. The unique contribution of this study is that it went beyond correlation analysis to identify an optimal cut-point and to test this cut-point. The study highlighted that further evidence is required regarding the ability of the CPT to accurately identify persons with and without impairment, in order to promote accurate test interpretation. Testing the CPT against a criterion related to outcomes in the community is needed.

Harm outcome study

The primary contribution of this study is that it tested the trustworthiness of an ecologically valid, real-world outcome. This study contributes to the thesis research by providing a foundation for a third study, examining the ability of functional measures to predict a harm outcome. The measurement of the incident of harm outcome in the community after discharge relied of necessity on proxy reporting, and the trustworthiness of the proxy reporting method required testing.

The study built on the method of Tierney and colleagues (2007) by using a more encompassing definition of “incident of harm” that included minor incidents of

harm that did not require medical attention. Thus, the study contributes by including incidents that may be considered by caregivers as less medically urgent, and proxy report of these incidents was reliable. It is important to include these incidents because, although perceived as medically minor, there may be persons admitted to long-term care after an accumulation of several minor incidents.

This study demonstrated that the data from this method was a valid estimation of the true occurrence of incidents of harm. It confirms other evidence demonstrating that proxy report for an observable outcome was trustworthy. The study identified advantages in using a telephone interview for data collection from proxy caregivers. Compared with written daily logs, the telephone interview had a higher response rate and allowed the interviewer to clarify information if needed. It also showed that it was difficult to rely on medical reports for detailed information about minor incidents of harm. Documentation of this method is useful for researchers interested in conducting studies to validate measures for ecologically valid real-world outcomes (Burgess, Alderman, Forbes, Costello, Coates, Dawson, et al. 2006). This outcome measurement method can be used to support further research related to incidents of harm. It is measured in the person's home setting and takes into account the real circumstances encountered in the home such as the environment or caregivers. The incident of harm outcome is based on a description of real consequences, rather than perception of safety or fear of future injury. As such, it can provide a useful outcome against which to examine the consequences and risks associated with living in the community.

Predictive validation study

The contribution of this study was to offer data to support the ability of in-hospital outcome measures to predict incidents of harm after discharge. These are the first data supporting the ability to predict harm prospectively for any older adult cognition measure in occupational therapy (Douglas, Letts & Liu, 2008). The support for this validity is important for three reasons: it asserts that an in-hospital occupational therapy functional assessment has meaning to real-world outcomes, it can promote the use of standardized measures, and it enables accurate test interpretation which in turn promotes improved care for patients.

An important finding was that two functional measures of cognition (AMPS-P and CPT) were stronger predictors of the incident of harm outcome than motor skills (AMPS-M), co-morbidities, burden of care (FIM) or demographic variables. It can be hypothesized that a person with frailty, poor balance or impaired fine motor skills may be able to compensate for these deficits in daily life, but poor cognitive skills make a person more susceptible to incidents of harm. This suggests a reduced emphasis on these weaker predictors of incidents of harm and a focus on the strongest predictors including cognitive skills, living alone, and age when considering future research and decision-making in clinical practice.

The manuscript provides data that offers interpretation of test scores to patients, families and the health care team with respect to incidents of harm. This is in addition to the interpretation already suggested by the measures with respect to level of independence. The scores are best interpreted alongside information about

living alone, age and sex. The results of the study caution against over-interpretation of scores, specifically, over confidence in determining impairment by the measures. However, they also help to mitigate under-interpretation. While over-interpretation of scores may happen more commonly in current clinical practice, an over cautious approach may lead to under-interpretation (e.g. assuming that performance in a specific task at a moment in time is relevant only to that task/time). This reduces the purpose of measurement to a narrow role of recording a singular observation, to the point one would question the purpose of clinical assessment. This study offers data that validates the measures for a specific purpose in a specific population. Given similar circumstances, the therapist can offer data that supports a link between the test score and harm outcomes.

Overall Implications

The thesis has implications for applying measurement theory to validation of measures in occupational therapy. The implications for use of the measures will be discussed first by demonstrating the validation argument for each measure, and ending with a comment about the usefulness of the measures in usual practice.

Validation of measures and measurement theory

The choice of prospective design and incident of harm outcome demonstrated a method for validation that is needed in applied health sciences such as occupational therapy. The rationale for the overall study design drew on the work

of Messick (1995) who argued that validation must address the real-world interpretation of a measure. He made his argument with emphasis on high-stakes testing. In this study, the consequences of testing versus not testing may not have differed significantly, but the consequences of inaccurate interpretation of observations and spurious recommendations may be significant. Families and older adults deciding about discharge home could potentially limit a person's autonomy based on recommendations of an occupational therapy assessment. Of measures used in occupational therapy studies, validation studies of measures have examined the ability to predict discharge location (Macneill, Gerskovich, Caron, Lichtenberg, 1997), living status (Loeb, 1996), or performance on another measure of community function (Velligan et al., 1998). The weakness of these outcomes is that the decision about discharge often is based on the test (criterion contamination) (Streiner & Norman, 2003). In this study, it was important to examine a post-discharge outcome that was not contaminated by the in-hospital assessment. Thus, the rationale for this study was based on the argument that validation needs to address the real-world interpretation of the measure and that assessment of the risk for harm was clinically needed.

Classical test theory informed the conception and analysis for the CPT study and Incident of harm study. For the CPT study, understanding of concurrent and construct validity highlighted the requirement for data supporting not only concurrent validity in general, but measures that tested whether the CPT is a measure of the construct of cognition independent of motor skills, age and

education. In the Incident of harm study the trustworthiness of the incident of harm outcome required testing using the concepts of test-retest reliability and validity of incident reporting. An understanding of classical test theory was important for ensuring the accuracy and completeness of the analysis. For example, in the “CPT study” analysis of cut scores allowed more precise interpretation than correlations alone.

The third manuscript (Predictive validation study) employed a line of argument in the study rationale based on the writings of Kane (2006). This theoretical conception of validation was examined during the writing phase of this third manuscript. Kane argues that validity is a function of purpose to which a measure is put. For the purpose of hospital discharge planning, it is noted that testing of the specific ability to predict harm is needed if we are to use measures to make recommendations about the risk for such harm. The thesis directly addressed the exact purpose for assessment, which is commonly to assess for safety or risk. Whether or not the measures were validated for this purpose, the thesis demonstrated a method for further testing of measures, which are used to predict the risk for harm in the community for persons with dementia.

Understanding the conception of argument-based validation as described by Kane (2006) reinforced that validation is based on sound research design, and that research must address the particular use of the measure. Argument-based validity theory guided the way the results were discussed, that is, the need to develop a link with other literature to demonstrate interpretive as well as empirical argument that

the measures would reasonably link to a harm outcome. Application of the theory placed the results of this study in context as part of a program of research towards validation of an instrument rather than declaring a measure valid based on a single study. As well, it clarified the conception of directions for future research.

Explanatory models

The purpose of this study was to examine the function of two measures and provide a predictive and therefore parsimonious model for harm outcome. The focus of the discussion was on the validity of the AMPS and CPT, however other independent variables had an effect on the outcome. Explanatory models for time to incident of harm would incorporate the variables living alone, age and sex. Furthermore, variables AMPS-motor score and care-giving hours were found to have moderate effect size for the occurrence of an incident of harm and would need to be included in the further study of explanatory models.

When considering the older adults at home and how these factors might interact to lead to harm, it is important to consider theory of aging. A theoretical model of aging such as the model of selection and optimization with compensation (SOC) (Freund & Baltes, 1998) can be employed to understand the older adults at home in this study. Older adults in this study who had a recent hospital admission were seeking to maintain the ability to manage their daily living tasks. The SOC model provides explanation for how persons adapt and maintain independence as they age. It states that both opportunities and limitations can be mastered

adaptively by manipulating three components: selection, optimization and compensation. Selection occurs when, due to limitations in resources, individuals must select which goals they want to pursue at any given time. For example in this study, participants may have selected not to use the stove at all to avoid safety hazards, and maintain their independence at home by using the microwave or meals on wheels services. Optimization is the allocation of resources so that the highest level of functioning can be achieved within the selected goals. For example, participants in this study may have put all their effort into bathing or hygiene to prevent a recurrent infection. Compensation is used when, faced with a decline in resources, individuals need to be able to identify substitutive processes that would allow them to maintain a certain level of functioning within their target domain. For example, a participant aware of declines in memory for taking medications, may have compensated with a blister pack or other reminders. SOC coping behaviours take on increased meaning in the old and very old adult when limitations in physical and cognitive resources and the use of compensatory mechanisms become more apparent (Lang, Rieckmann, & Baltes, 2002).

In this study, the variable “living alone” had significant effect on whether or not harm occurred. This result can be interpreted to indicate that persons who lived alone were not able to benefit from a protective effect. Persons who did not live alone may have had the benefit of a caregiver to check on them and to prevent harm from occurring or ensure that that “near misses” did not become events that led to injury. However, with consideration of the SOC model of aging, the link

between living alone and harm outcome may be more complex. Persons who did not live alone may have had greater resources upon which to draw upon when selecting, optimizing or compensating for deficits. They may also have had another person to aid in the generation of strategies to select, optimize or compensate. Given that care-giving hours may have had moderate effect on the occurrence of harm, it may be that care-giving enabled compensatory strategies in this sample. However, what makes it complex is that persons who lived alone may have been able to live alone simply because they had greater capacities, including greater abilities to use SOC strategies. Persons in this sample differed from the larger population in that they had experienced a serious medical illness that had required hospitalization and rehabilitation. Therefore, the find that living alone was linked with increased harm outcome requires replication in a larger population, using measures of minor and major injury to determine whether living alone is a risk factor for various types of harm in the short term (0-6 months) or long term.

Clinical interpretation of each functional measure

The study can contribute to improved care for persons with dementia and to occupational therapy because it offers evidence to support the appropriateness of using occupational therapy functional measures to fill the need for pre-discharge measures. Findings that the AMPS and CPT predicted incident of harm outcome contributes to the body of research on the AMPS and CPT by supporting what Kane (2006) calls the interpretive argument. The interpretive argument for the validation

of these measures in prediction of harm includes the rationale for why the measure is expected to relate to the outcome, including the theoretical construction of the tool, and further testing of the measure. Further testing can include several concepts and the focus of this study was to test extrapolation and decision-rules.

For each measure, the validation argument (Kane, 2006) will be summarized and made explicit in the following paragraphs. It was anticipated that both the measures are related to harm because (a) other cognitive screening measures are reported to be related to harm (Tierney et al., 2007), (b) previous validity data on the measures revealed an association between both CPT scores and time to long-term care admission (Burns, Mortimer & Merchak, 1994) and between AMPS-process scores and another checklist for safety in the home (McNulty & Fisher, 2001), and (c) most importantly, because several studies show a link between declines in cognition and measures of morbidity (Banaszak-Holl et al., 2004; Tun, Murman, Long, Colenda, & von Eye, 2007).

To demonstrate the validation argument for each measure, the theoretical construction and testing of both extrapolation and decision rules is described. In this way, the coherence of the argument for validation is shown (Kane, 2006). The AMPS is constructed based on task analysis using objective observation of individual components during task completion. The scoring of the AMPS- Process scale requires detailed recording of specific observable behaviours rather than drawing inferences about various cognitive constructs. As well, the AMPS was developed with the assumption that assessment of routine tasks, chosen by the participant,

provide the most accurate information about actual performance in daily life. It is presumed that the participant can thus optimally perform in the testing situation, reflecting true skill more accurately and with less influence of education skills. Previous extrapolation testing includes data previously cited regarding reliability, concurrent validity and ability to discriminate between groups. This current study provided further extrapolation data demonstrating that the AMPS-Process scale was a predictor of incidents of harm. Decision rule testing, or sensitivity and specificity of the tool (Kane 2006) was undertaken. The AMPS-Process scale cut-scores had low sensitivity and specificity for identifying a person who had any incident of harm over six months. Application of argument based validity therefore demonstrated validation of the AMPS-process scale for predicting harm in older adults six months after discharge, but showed, when using the scale for this purpose, it had limited precision. Thus, there is a coherent argument for validity of the AMPS for this purpose, with clearly stated limitations in test cut-score interpretation. The validation argument therefore relies on the correlation with increased risk for harm and demonstrates the limitations in the use of cut-score for decision-making.

The CPT theoretical construction is based on an understanding of working memory (Burns, 1992). The purpose of the test is to observe performance of set tasks, which were designed to examine common daily living skills. Scoring is based on the amount of cueing required by the participant and is interpreted to indicate working memory ability. The CPT study showed that scores are related to other cognitive measures but the decision rules are not similar in that it does not identify

persons who are “independent” consistently with other measures. The predictive validation study provided further extrapolation data that demonstrated moderate predictive ability of the CPT for incidents of harm in persons with dementia. The CPT showed high specificity for identifying persons without any incident of harm in this sample. The validity argument for the CPT shows coherence but again, there are significant caveats for interpretation of the scores. The CPT demonstrates a more coherent validation argument when used for the purpose of identifying persons with dementia who are at lowest risk for an incident of harm.

The description of the validation argument therefore demonstrates the overall coherence of the argument for validation of the AMPS and CPT for the explicit purpose of predicting harm in older adults six months after hospital discharge. The argument demonstrates the limitations of validation for the harm outcome, but does not invalidate the measures for the purpose for which they were originally intended, which was to identify the level of independence. This study contributes by adding evidence to each measure’s overall program of validation.

Lastly, an overall implication of this study is that it illustrates the clinical usefulness, including feasibility and acceptability of the AMPS and CPT in a rehabilitation hospital. Both measures were administered on the hospital unit during care. They were tolerated by the study participants, as none withdrew during or as a result of assessment administration. Currently, on the hospital unit the occupational therapists use either assessment with the patients referred (feasibility). Although in some settings, time would not allow administration of both

measures, or self-report of caregivers, the more complex nature of the caseload in this setting promotes direct observation using standardized assessments. The clinicians report that both measures are acceptable to the patients because they employ daily tasks that are easily understood by the patients and families to relate to daily living upon discharge. Both measures are used, however the AMPS is used on admission and discharge to measure change. Clinicians reported using the CPT because the test interpretation provided in the manual includes tips for caregivers on how to care for persons with dementia. By demonstrating the value of these tests in predicting harm outcomes, the preliminary study results encouraged managers to support staff in obtaining and using the measures as part of the services offered by the program. This demonstrates feasibility and acceptability in a particular setting, and adds to the overall implications for the future clinical impact of the study results.

Limitations

Limitations specific to each study have been noted in the individual papers, but some overall limitations of the thesis as a whole are more broadly discussed here. Study limitations will be described in terms of feasibility limitations and ethical considerations. In addition to the baseline measures (e.g. demographic information, SMMSE, FIM, co-morbidities), other patient characteristics were not tracked such as caregiver stress, home environment characteristics (grab bars, stairs) and uptake of rehabilitation intervention (e.g. correct use of ambulation aids). These factors may

have influenced the harm outcome more than the in-hospital score on a functional test. Study feasibility influenced the length of follow-up period to six-months. It is not known how this influenced the correlations found with the baseline measures, whether stronger or weaker correlations may have been found. Study timelines and budget also limited the sample size enrolled, which limited the regression modeling to include two variables.

Study recruitment was limited by ethical considerations. Both caregiver proxy participants and patient participants were required to provide consent. It is possible that the demographics of the sample may have differed with recruitment of only one or other of the participants in the pair. Persons who did not have a consenting caregiver proxy may have had shorter time to incident of harm than those in this sample. As well, persons with dementia who did not consent may have had depression, paranoia, or behavioural issues related to dementia and may have experienced harm outcomes sooner than persons in this study.

The design of the overall study was such that testing of reliability and validity of the harm outcome measure was simultaneous, and completed on the same sample. This places limits on the reproducibility of reliability and validity of the harm outcome. For example, return rates for the caregiver daily logs were lower than anticipated, and may have shown lower rates of agreement with higher return rates. Ideally the harm outcome would have been tested first to confirm the optimal administration of the outcome measure for the predictive validation study.

The definition of safety as “incidents of harm” was narrow due to the need for observability and objectivity of the outcome. One could conceive of safety in a broader sense, as a perception of risk and sense of security (Parmelee & Lawton, 1990) rather than observed incidents. The caregiver interview was very specific to ask about incidents. A conception of safety that included perception of risk may have taken into account major versus minor incidents rather than the time or frequency. A major incident of harm such as a house fire or fracture from a fall may have greater sway on decision-making for long term care. Perception of risk in the future also may affect decision-making, for example several persons enrolled in the study moved to assisted living even though no incident of harm occurred, which was likely based on perceived risk of future events. A future study that measures both perception of risk and actual incidents simultaneously would be valuable for examining the relation of risk perception to actual risk, and explore the contributions of ageism, convenience, and preservation of autonomy.

Within these limitations, these studies offer the first information to link occupational therapy measures of function with a harm outcome and the methods used can inform future studies.

Future Research

Although the argument seems coherent that testing everyday tasks is a valid measure of risk for harm, this line of argument may be subject to validity fallacy (Kane, 2006). Given the findings that there was association with harm but limited

precision for the decision rules, the question arises whether stronger association can be found. This can be examined on two levels. First, there is the level of the constructs: cognition and harm. Secondly, there is the level of the measures: the ability of the CPT and AMPS to measure cognition and the ability of the Incident of Harm Caregiver Questionnaire to measure harm.

Examining the level of the construct involves examining the relationship between the variables of cognition and harm. An association was found in this study that was mediated by other variables. Further development of multivariate models for prediction of harm outcomes would have significant value for clinicians working with a population of older adults with dementia. Models including additional variables not tested in this study such as the physical home set-up, caregiver attributes (e.g. perceived burden) may have additional predictive value. A model containing the most predictive variables can be examined for sensitivity and specificity for identification of persons with an incident of harm. Such a model could be employed as a checklist for discharge planning, and factors that are modifiable could be made a priority in rehabilitation intervention programs. Secondary data analyses from longitudinal studies that include collection of data on injuries at home, such as the Canadian Longitudinal Study on Aging (Raina et al., 2009), may be of value for answering this research question. Data on the correlation between injuries and cognitive scores may corroborate these results, and regression models including demographic, ADL and motor skill variables may aid in determining the most predictive variables. Further data gathering on more detailed types of

unintentional harm such as failure to take medication correctly, or failure to eat, which are not currently being gathered by the Canadian Longitudinal Study on Aging would be invaluable.

At the construct level, questions arise about whether a specific cognitive component is related to harm. Everyday routine tasks may correlate with the harm outcome, but it may in fact be the ability to adapt to or to cope with novel situations that truly influences the harm outcome. The ability to complete novel tasks may be captured with measures of executive function such as the Executive Interview (Royall, Mahurin, & Gray 1992) or screening tools including executive function such as the Montreal Cognitive Assessment (Nasreddine, Phillips, Bédirian, Charbonneau, Whitehead, et al., 2005).

At the measurement level, the CPT and AMPS, in this study, may have actually measured the ability to deal with novel situations. Testing occurred in a hospital unit, rather than a familiar home environment. Indeed, a previous study found that persons with Alzheimer's disease scored higher on the AMPS in the familiar environment at home than in a clinic (Nygård, Bernspång, Fisher & Winblad, 1994). In contrast, persons with recent acquired brain injury scored higher in the hospital than home (Darragh, Sample & Fisher, 1998). Future studies are needed to compare the ability to predict measurement of harm in familiar versus unfamiliar environments using familiar versus novel tasks.

At the measurement level, further development of the Incident of Harm Caregiver Questionnaire is needed if it is to be used in future validation studies.

Additional testing of reliability and validity is needed to replicate the current results. The questionnaire can be an important criterion against which to validate other functional measures that show promise for prediction of harm outcomes for persons with dementia. The incident of harm outcome could serve as a criterion for validation of the AMPS or CPT in other populations (e.g. persons with acquired brain injury). As well, the number and frequency of incidents obtained by using the questionnaire may contribute an additional criterion against which to validate other measures.

Pragmatic consideration of funding and time restrictions on research, and the urgent need to test measures for real world outcomes lead to certain research priorities. Numerous tools are currently available for measurement of cognition (Douglas, Letts & Liu, 2008), which lessens the priority for the development of new tools. Given that the health care environment presses clinicians to assess efficiently and make recommendations about risk after discharge, the priority for further research is to further test the incident of harm outcome and examine which variables, including cognition and living alone, best predict harm outcome. This could ultimately result in identifying which factors may be modifiable and could be addressed as part of therapy interventions when helping older adults to plan for discharge to the community.

Conclusion

The results of this study have relevance to occupational therapy clinicians who commonly assess older adults and help to plan for discharge from the hospital. The study showed that two assessments used to examine daily tasks were predictive of incident of harm outcome after discharge in this sample. It also shows the limitations in using the measures as sole predictors of incidents of harm, and that other factors such as living alone must be taken into account. When using the measures to identify impairment, the lack of consistency between the two measures calls into question the precision of the identification of impairment. Therapists can communicate the evidence that the measures are predictive of harm after discharge, but the predictive ability has limited precision. This study supports the value of the occupational therapy assessments as part of the discharge planning process, and can form a foundation for further research on the most effective predictors for harm after discharge. This research is needed to help guide decisions about who is most at risk after discharge and would most benefit from scarce home care resources and rehabilitation support.

References

- Banaszak-Holl, J., Fendrick, A. M., Foster, N. L., Herzog, A. R., Kabeto, M. U., Kent, D. M., et al. (2004). Predicting nursing home admission: Estimates from a 7-year follow-up of a nationally representative sample of older Americans. *Alzheimer Disease & Associated Disorders, 18*, 83-89.
- Burgess, P. W., Alderman, N., Forbes, C., Costello, A., Coates, L., Dawson, D. R., et al. (2006). The case for the development and use of “ecologically valid” measures of executive function in experimental and clinical neuropsychology. *Journal of the International Neuropsychological Society, 12*, 194–209.
- Burns, T. (1992). The Cognitive Performance Test: An approach to cognitive level assessment in Alzheimer disease. In C.K. Allen, C.A. Earhart & T. Blue (Eds.), *Occupational therapy treatment goals for the physically and cognitively disabled* (pp. 46–84). Rockville, MD: The American Occupational Therapy Association, Inc.
- Burns T., Mortimer J.A., & Merchak, P. (1994). Cognitive Performance Test: A new approach to functional assessment in Alzheimer's disease. *Journal of Geriatric Psychiatry and Neurology, 7*, 46-54.
- Darragh, A. R., Sample, P. L., & Fisher, A. G. (1998). Environment effect on functional task performance in adults with acquired brain injuries: Use of the Assessment of Motor and Process Skills. *Archives of Physical Medicine and Rehabilitation, 79*, 418-423.
- Freund, A M., & Baltes, P. B. (1998). Selection, optimization, and compensation as strategies of life management: Correlations with subjective indicators of successful aging. *Psychology and Aging, 13*, 531-543.
- Kane, M. T. 2006. Validation. In R. Brennan (Ed.) *Educational Measurement, 4th edition*. Wesport, CT: American Council on Education and Praeger, 17 - 64.
- Lang, F. R., Rieckmann, N., & Baltes, M. M. (2002). Adapting to aging losses: Do resources facilitate strategies of selection, compensation, and optimization in everyday functioning? *Journal of Gerontology: Psychological Sciences, 57B*, 501-509.
- Loeb, P. A. (1996). *Independent Living Scales Manual*. San Antonio, TX: The Psychological Corporation.

- Messick, S. (1995). Validity of psychological assessment. *American Psychologist* 50, 741-749.
- MacNeill, S. E., Gerskovich, T., Caron, J., & Lichtenberg, P. A. (1997). Living alone: Predictors of recovery during medical rehabilitation. *Clinical Gerontologist*, 18, 3-13.
- McNulty, M. C., & Fisher, A. G. (2001). Validity of using the Assessment of Motor and Process Skills to estimate overall home safety in persons with psychiatric conditions. *American Journal of Occupational Therapy*, 55, 649-655
- Nasreddine, Z.S., Phillips, N.A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J.L., Chertkow, H. (2005). The Montreal Cognitive Assessment (MoCA): A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 695-699.
- Nygård, L., Bernspång, B., Fisher, A. G., & Winblad, B. (1994). Comparing motor and process ability of persons with suspected dementia in home and clinic settings. *American Journal of Occupational Therapy*, 48, 689-696.
- Parmelee, P. A. & Lawton, M. P. (1990). The design of special environments for the aged. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the Psychology of Aging* (3rd ed., pp. 464-488). New York: Academic Press Inc.
- Raina, P.S., Wolfson, C., Kirland, S.A., Griffith L.E., Oremus, M., Patterson, C. et al. (2009). Canadian Longitudinal Study on Aging (CLSA). *Canadian Journal on Aging* 28, 221-229.
- Royall, D. R., Mahurin, R. K., & Gray, K. F. (1992). Bedside assessment of executive cognitive impairment: The executive interview. *Journal of the American Geriatrics Society*, 40, 1221-1226.
- Streiner, D. L., & Norman, G. R. (2003). *Health measurement scales: A practical guide to their development and use* (3rd ed.). New York: Oxford University Press.
- Tierney, M. C., Snow, W. G., Charles, J., Moineddin, R., & Kiss, A. (2007). Neuropsychological predictors of self-neglect in cognitively impaired older people who live alone. *American Journal of Geriatric Psychiatry*, 15, 140-148.
- Tun, S. M., Murman, D. L., Long, H. L., Colenda, C. C., & von Eye, A. (2007). Predictive validity of neuropsychiatric subgroups on nursing home placement and survival in patients with Alzheimer disease. *American Journal of Geriatric Psychiatry*, 15, 314-327.

Velligan, D. I., Bow-Thomas, C. C., Mahurin, R., Miller, A., Dassori, A., & Erdely, F. (1999). Concurrent and predictive validity of the Allen Cognitive Levels Assessment. *Psychiatry Research, 80*(3), 287-298.

PhD thesis – A. Douglas
McMaster University – Rehabilitation Science

Appendix 1
Research Ethics Board final approval letter



RESEARCH ETHICS BOARD



REB Office, 1057 Main St. W., Hamilton, ON L8S 1B7
Telephone: 905-521-2100, Ext. 42013
Fax: 905-577-8379

Research Ethics Board Membership
Jack Holland MD FRCP FRCP(C)
Chair
Suzette Salama PhD
Vice-Chair/Ethics Representative
Mary Bedek CCHRA (C)
Privacy Officer
Morris Blajchman MD FRCP(C)
Hematology
Julie Carruthers MLT
Research, Transfusion Medicine
Adriana Carvalhal MD, MSc, PhD
Psychiatry
David Clark MD PhD FRCP(C)
Medicine
Jean Crowe MHSc
Rehabilitation Science
Lynn Donohue BA(Hons)
Community Representative
Brock Easterbrook BA
Research Coordinator, Anaesthesia
Farough Farokhyar Mphil, PhD, Pdoc
Surgery
Sylvia Fung BSP LLB
Pharmacy/Legal
Melanie Griffiths FRCP (UK)
Diagnostic Imaging
Cindy James BScN
Gastroenterology
David Jewell M.S.W, MHSC
Geriatrics
Norman Jones MD FRCP FRCP(C)
Medicine
Peter Kavsak PhD
Laboratory Medicine
Rosanne Kent RN BA MHSc(M)
Cardiology
Carolyn Kezel, RN
Pediatrics
Shelly McLean, MBA
Community Representative
Madhu Natarajan MD, FRCP, FACC
Cardiology
Kesava Reddy MB BS FRCS FACS
Neurosurgery
Susan Rivers RN MSC(T)
Geriatrics
Gita Sobhi BSc Pharm
Pharmacy
Brian Timmons, PhD
Pediatrics
Marie Townsend BA(Hons), MBA
Administration
Graham Turpie MD FRCP(C)
Medicine
Jeff Weitz MD FRCP(C) FACP
Medicine
Jim Wright BSc MD
Radiation Oncology
Ed Younglai PhD
Obstetrics/Gynecology

The HHS/FTS REB operates in compliance with the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans; the Health Canada / ICH Good Clinical Practice: Consolidated Guidelines (E6); and the applicable laws and regulations of Ontario. The membership of this REB also complies with the membership requirements for REBs as defined in Canada's Food and Drug Regulations (Division 5: Drugs for Clinical Trials Involving Humans Subjects).

October 27, 2008

PROJECT NUMBER: 08-424
PROJECT TITLE: Home After Hospital Occupational Therapy Assessment
PRINCIPAL INVESTIGATOR: Dr. Lori Letts

This will acknowledge receipt of your letter dated October 20, 2008 which enclosed a copy of the revised Information Sheet/Consent Forms for the above-named study along with a response to the specific issues raised by the Research Ethics Board at their meeting held on September 16, 2008. Based on this additional information, we wish to advise your study has been given **final** approval from the full REB. The submission, revised Protocol version dated October 20, 2008, including the Information Sheets/Consent Forms—Contact Person and Patient was found to be acceptable on both ethical and scientific grounds. **Please note** attached you will find the Information Sheet with the REB approval affixed; all consent forms used in this study must be copies of the attached materials.

We are pleased to issue final approval for the above-named study for a period of 12 months from the date of the REB meeting on September 16, 2008. Continuation beyond that date will require further review and renewal of REB approval. Any changes or amendments to the protocol or information sheet must be approved by the Research Ethics Board.

The Hamilton Health Sciences/McMaster Health Sciences Research Ethics Board operates in compliance with the ICH Guidelines Good Clinical Practice Guidelines, the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans and Division 5 Health Canada Food and Drug Regulations.

Investigators in the Project should be aware that they are responsible for ensuring that a complete consent form is inserted in the patient's health record. In the case of invasive or otherwise risky research, the investigator might consider the advisability of keeping personal copies.

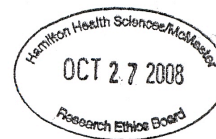
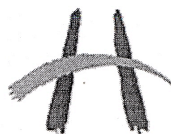
A condition of approval is that the physician most responsible for the care of the patient is informed that the patient has agreed to enter the study.

PLEASE QUOTE THE ABOVE-REFERENCE PROJECT NUMBER ON ALL FUTURE CORRESPONDENCE

Sincerely,

Jack Holland, MD, FRCP, FRCP (C)
Chair, Research Ethics Board

Appendix 2
Information letter and consent for patient participants



Health Sciences **Hamilton Health Sciences**
RESEARCH INFORMATION SHEET

Title of Study: Home after Hospital Study

Locally Responsible Investigator: Lori Letts, PhD. Associate Professor
School of Rehabilitation Sciences, McMaster University

Principal Investigator (student): Alison Douglas, MScOT, PhD student,
School of Rehabilitation Sciences, McMaster University

You are invited to participate in a research study because you are on the Complex Medical Rehabilitation Unit. This form gives you information so you can decide if you want to participate or not. If you want to join, you will be asked to sign a form. Please take your time to make your decision. Feel free to talk about it with your family, friends or doctor.

WHY ARE WE DOING THIS STUDY?

Older adults often return home after being in the hospital. When in hospital, they can be seen by an occupational therapist (OT) to manage daily living (like dressing or cooking). Research is needed to know how different OT assessments predict if people manage safely at home. The study is needed to make sure we are doing the best we can to know how you will manage at home. About 40 people will be involved in the study. This information will improve occupational therapy. Also, it will help other older adults and their families make decisions about how they will manage safely at home.

WHAT BENEFITS ARE INVOLVED IN THIS STUDY?

A person who knows you will get a monthly phone call and general information about where to receive help if needed.

WHAT RISKS ARE INVOLVED?

We are not aware of any risks to being in this study, but you should know that we will be talking with a caregiver about how you are doing with daily living after you leave the hospital.

PhD thesis – A. Douglas
McMaster University – Rehabilitation Science

WHAT DO I NEED TO DO?

If you volunteer to participate in this study, we will ask you to do the following:

- To accept monthly telephone calls to get information about the services the participant in this study receives and any changes in his/her health or injuries over the next 6 months. You will only be contacted if the participant consents to this study.
- To complete a diary checklist about any changes in the participants health for one month of the 6-month study period

WHAT ELSE DO I NEED TO KNOW?

You have no obligation to be in the study. The care and treatment of the participant will be the same if you are in the study or not. Even if you agree to be in this study, you can withdraw at any time. You can also remove your data from the study if you want.

Your information will not be shared with anyone except with your consent or as required by law. Your information will be confidential and locked in an office. If the results of the study are published, your name or anything to identify you will not be published. This signed consent form, may be included in your hospital chart.

If you have any questions about the research now or later, please contact Alison Douglas ((905) 525-9140 X 21482). If you have any questions regarding your rights as a research participant, you may contact the Office of the Chair of the Hamilton Health Sciences/Faculty of Health Sciences Research Ethics Board at 905-521-2100, ext. 42013.



Informed Consent Form

Title of Study: Home after Hospital Study

1) Yes, please sign me up. I hereby consent to partake in this study.

Name of Participant (Print) Signature of Participant Date

Name of Substitute decision maker Signature of Substitute decision maker Date
(if required)

2) I agree that my family doctor will be contacted to get health information about me while I am in this study.

Signature of Participant Date

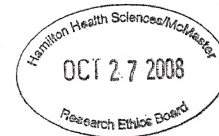
3) I agree that a contact person will be phoned to get information, including the services I receive and any changes in my health or injuries while I am in this study.

Signature of Participant Date

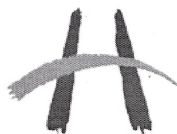
Name of caregiver participant: _____
Relationship to participant: _____
Contact information: phone: _____

Person obtaining consent (print) Signature of person obtaining consent Date

Principal Investigator (print) Signature of Principal Investigator Date



Appendix 3
Information letter and consent for caregiver participants



Health Sciences **Hamilton Health Sciences**
RESEARCH INFORMATION SHEET: CONTACT PERSON

Title of Study: Home after Hospital Study

Locally Responsible Investigator: Lori Letts, PhD. Associate Professor
School of Rehabilitation Sciences, McMaster University

Principal Investigator (student): Alison Douglas, MScOT, PhD student,
School of Rehabilitation Sciences, McMaster University

You are invited to participate in a research study because you give care to the participant who is on the Complex Medical Rehabilitation Unit. This form gives you information so you can decide if you want to participate or not. If you want to join, you will be asked to sign a form. Please take your time to make your decision.

WHY ARE WE DOING THIS STUDY?

Older adults often return home after being in the hospital. When in hospital, they can be seen by an occupational therapist (OT) to manage daily living (like dressing or cooking). Research is needed to know how different OT assessments predict if people manage safely at home. The study is needed to make sure we are doing the best we can to know how you will manage at home. About 40 people will be involved in the study. This information will improve occupational therapy. Also, it will help other older adults and their families make decisions about how they will manage safely at home.

WHAT BENEFITS ARE INVOLVED IN THIS STUDY?

A person who knows you will get a monthly phone call and general information about where to receive help if needed.

WHAT RISKS ARE INVOLVED?

We are not aware of any risks to being in this study but you should know that we will be asking you about how the participant is doing with daily living after he/she has left the hospital.

PhD thesis – A. Douglas
McMaster University – Rehabilitation Science

WHAT DO I NEED TO DO?

If you volunteer to participate in this study, we will ask you to do the following:

- To accept monthly telephone calls to get information about the services the participant in this study receives and any changes in his/her health or injuries over the next 6 months. You will only be contacted if the participant consents to this study.
- To complete a diary checklist about any changes in the participants health for one month of the 6-month study period

WHAT ELSE DO I NEED TO KNOW?

You have no obligation to be in the study. The care and treatment of the participant will be the same if you are in the study or not. Even if you agree to be in this study, you can withdraw at any time. You can also remove your data from the study if you want.

Your information will not be shared with anyone except with your consent or as required by law. Your information will be confidential and locked in an office. If the results of the study are published, your name or anything to identify you will not be published. This signed consent form, may be included in your hospital chart.

If you have any questions about the research now or later, please contact Alison Douglas ((905) 525-9140 X 21482). If you have any questions regarding your rights as a research participant, you may contact the Office of the Chair of the Hamilton Health Sciences/Faculty of Health Sciences Research Ethics Board at 905-521-2100, ext. 42013.



Informed Consent Form

Title of Study: Home after Hospital Study
Consent from contact person

Yes, please sign me up.

I agree that I will be contacted by telephone during this study (6 months) to give information about the participant including services received and any changes in health or injuries.

Name of Participant (Print)

Signature of Participant

Date

Person obtaining consent (print)

Signature of person obtaining consent

Date

Principal Investigator (print)

Signature of Principal Investigator

Date

