RELIABILITY AND VALIDITY TESTING OF AN ACUITY TOOL FOR USE IN THE LONG-TERM ACUTE CARE HOSPITAL SETTING

A DISSERTATION

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

IN THE GRADUATE SCHOOL OF THE

TEXAS WOMAN'S UNIVERSITY

COLLEGE OF NURSING

BY

KATHRYN S. SPIEGEL, M.S.

DENTON, TEXAS

MAY 2001

ACKNOWLEDGMENTS

I would like to express my sincere appreciation to the members of my committee for all of their support and assistance with this dissertation: Dr. Lynn Wieck, who kept me on track through her positive "you can do it" attitude and provided me with excellent editorial assistance; Dr. Ann Malecha, who, without her creative thinking, this project may have never come to fruition; Dr. Miguel daCunha, who has been a support system for a previous as well as the current (and last) academic endeavor; and Dr. Anne Young who never doubted that this could be done.

Thanks go to my sons, David and Scott, and my husband, Joe, for their support and encouragement. Finally, a special thank you goes to Verna Deinla, RN, the messenger.

RELIABILITY AND VALIDITY TESTING OF AN ACUITY TOOL FOR USE IN THE LONG-TERM ACUTE CARE HOSPITAL SETTING

ABSTRACT

KATHRYN S. SPIEGEL, M.S.

TEXAS WOMAN'S UNIVERSITY
COLLEGE OF NURSING
MAY 2001

The new millennium will see increased health care costs coupled with increased need for services--all compounded by a nursing shortage. In addition, regulatory agencies are mandating that standards be used to benchmark the quality of care provided to today's consumer. Future success of health care organizations will be measured by their ability to effectively manage resources to provide quality care at the appropriate level with competent personnel.

Classification systems have been in existence for over 30 years. Different classification systems have been developed to meet the individual patient characteristics in a variety of hospital settings. The uniqueness of patient care needs prohibits generic classification systems. Formatting a system according to the characteristics of the patient population in which it will be used is essential.

The purpose of this study was to determine psychometric properties of an acuity tool for use in the long-term acute care population. Internal consistency reliability was tested with Cronbach's alpha; an expert panel using the diagnostic content validation model (DCV) determined content validity. Two hundred acuity tools from four patient care areas of a long-term acute care hospital were used for psychometric testing.

A one-way ANOVA indicated significant differences in acuity scores by location. The ICU had the highest overall scores, followed by the IMU and general medical-surgical areas. Differences were noted in number of activity points, which suggested that higher acuity is directly proportional to intensity of nursing care.

A Cronbach's alpha coefficient of 0.7538 indicated that the tool was a reliable measure of acuity in long-term acute care population. The DCV coefficient was 0.84, which supported the hypothesis that the instrument was a valid tool for evaluating acuity in this population. Factor analysis yielded four factor trends among the variables. Each subscale on the LTA-Care Acuity Index loaded on one or more factors with a minimum coefficient of 0.30.

This study's findings have demonstrated that reliability and validity of an acuity instrument can be tested to provide hospital personnel, regulatory agencies, and health care consumers with a system for measuring acuity in long-term care facilities.

TABLE OF CONTENTS

Chapter

APPENDIXES

LIST OF TABLES

Table

LIST OF FIGURES

Figure

CHAPTER 1

INTRODUCTION

Patient classification systems based upon assessed patient care needs are being developed or refined in many types of patient care settings. These systems are designed to predict and measure patient care needs as they relate to staffing requirements, scheduling, and budgetary allocations (Giovanetti, 1979). The prospective payment system (PPS), health maintenance organizations (HMOs), and changes in Medicare reimbursement have had an additional impact in terms of urgency for the initiation and use of patient classification systems for allocating nursing care resources.

Patient classification systems have been in use in the acute care hospital setting for over a quarter of a century. Many different types of classification systems along with patient care delivery models have been used to manage nursing resources.

There are four major types of patient classification systems: (a) disease, (b) procedure, (c) acuity, or (d) some combination of the three (Bermas & Van Slyck, 1984). Of these four classification methods, the acuity-based system is the most frequently used system in hospital nursing departments. Acuity systems measure the patient's nursing care requirements based on a need for bedside therapeutic nursing interventions as well as the patient's psychosocial dependency levels. Acuity systems

provide a guide for predicting staffing requirements, scheduling personnel, and providing support for annual budget allocations (Whitney & Killien, 1987).

One of the most widely used scoring systems for measuring patient acuity is the Therapeutic Intervention Scoring System (TISS). Developed in 1974, the TISS measures severity of illness by quantifying the therapeutic interventions applied to a patient (Cullen, Civetta, Briggs, & Ferara, 1974). The acuity system has enabled nursing administrators to gain insight into the intensive care unit's (ICU's) severity of illness, utilization of ICU facilities, required ICU nurse/patient ratios, and the number of hospital beds needed for ICU purposes. Recently, two new versions of the TISS have been published: (a) the Intermediate TISS (Cullen, Nemeskal, & Zaslavsky, 1994), which adapts the use of the instrument to non-ICU patients; and (b) the TISS-28, which is a simplified version of the 85-item Intermediate TISS (Reis Miranda, De Rijk, & Schaufeli, 1996). All three instruments use the same scoring pattern for each activity with points ranging from 1 to 4. Higher points indicate more complex, time-consuming interventions.

Long-term acute care hospitals have been in existence for approximately 15 years. They were developed as transition models for patients requiring extended care in a highly skilled environment. Long-term care facilities are diagnostic related group (DRG) exempt; that is, there is not a limitation in terms of length of stay for a particular illness or diagnosis. Nursing care requirements representing tasks and interventions are similar to those of a short-term acute care setting. Differences lie in

the types of patient care activities performed by the nurse including ventilator management, enteral feedings, wound care, patient immobility, and psychological support. In the acute care setting, these elements of care may be limited to the ICU patient where the nurse/patient ratio does not exceed one to three. In the long-term care setting, these tasks are performed in a general medical-surgical unit, where a nurse may be responsible for six to seven patients.

Instrument development to measure physiological and psychosocial phenomena is a process that begins with defining and analyzing a concept, writing questionnaire items, and choosing the scaling and scoring methods (Summers, 1993). The Long-Term Acute Care Acuity Index (LTA-Care) (Appendix A) is patterned after the point structure format of the TISS. However, it differs in several ways. The instrument is divided by body systems to include pulmonary, cardiovascular, gastrointestinal, renal/genitourinary, integumentary, and neurological systems. Additional categories include: psychosocial interventions, special procedures, daily activities, and intravenous lines, including blood product administration and blood glucose monitoring. The nursing time involved in medication administration is accounted for by adding the total number of medications given in a 24-hour period to the total intervention points. The design format permits multiple administrations of the instrument to the same patient and allows for trending of acuity scores as patients' conditions change and/or patients move to different levels of care.

Problem of Study

The purpose of this study was to examine the reliability and validity of the Long-Term Acute Care (LTA-Care) Acuity Index as an objective measurement of acuity in long-term, acute care, hospitalized patients. Each item contained in the acuity tool was tested individually and collectively as indicators of acuity in this patient population.

Rationale for Study

Patient classification systems allow for efficient and effective management of nursing resources. The management of these resources will aid in decision-making regarding care priorities. Resource management can result in improved nursing care and appropriate utilization of services. All of these factors can contribute to both job satisfaction for the providers of patient care as well as patient satisfaction with the delivery of care.

Hospitals are seeking ways to identify costs and eliminate unnecessary expenses. Because nursing costs generally compose the largest portion of the personnel budget, they must be measured and justified to avoid unwarranted reductions (Ebner, 1985). It was estimated that hospitals spend $15,000,000 each year on nursing staffing studies (Vaughn & MacLeod, 1980). The objective of these studies was to establish a system that, according to the number of patients and their care

needs, balances the nursing staff available on each unit for each shift with the manpower required (Jackson & Resnick, 1982).

Reliability and validity testing are essential when acuity systems are used in staffing patterns, budgeting, and predicting the cost of nursing services. Reliability and validity testing answers the following questions: (a) to what degree does the instrument measure actual requirements of nursing care to hospitalized patients; (b) is the tool a reliable and valid instrument for the selected patient population; and (c) are patients who require the same level of care classified with similar acuity scores? Measurement of reliability and validity precede instrument utilization when: (a) an established instrument is used in a new setting; (b) an existing instrument is modified for use in a new setting; or, (c) a new instrument is developed for the same purpose or to measure the same construct in a new setting.

Theoretical Framework

Measurement consists of the use of rules that quantify attributes of people or objects so that these attributes can be described or explained (Nunnally, 1970). Psychometric theory is concerned with the issues and methodologies related to the basic principles of measurement. Classical measurement theory, also known as the theory of measurement error or the true-score model, is a psychometric theory that was appropriate in this study.

The basic principle of classical measurement theory is that the observed (obtained) score is the sum of the true score and the error score (Nunnally, 1959, 1967). It is apparent that increased measurement error will result in decreased reliability and validity of the measurement. The effect of random error may also be negated and the true score more accurately assessed by determining the average of many independent measures (Nunnally, 1959).

Nunnally (1970) identified several possible sources of measurement error. These include poorly standardized instructions, errors in scoring, subjectivity of measurement, variation, variation in testing conditions, errors due to guessing, item or subject sampling bias, and transitory personal factors. Decreasing these sources of error will increase the reliability and validity of the measure (Nunnally, 1970).

There are two major frameworks for measurement: (a) criterion-referenced and (b) norm-referenced. Classical measurement theory forms the basis for norm-referenced measures with the premise that every score includes some random error (Waltz, Strickland, & Lenz, 1991). Norm-referenced measurement depicts a range of possible scores and evaluates the performance of subjects against each other along a continuum reflective of a normal dispersion. Norm-referenced tests discriminate among subjects in the sample by reflecting the amount of an attribute that subjects possess. Variance is the essential feature of norm-referenced measures. The LTA-Care Acuity Index is a norm-referenced measurement tool.

Measurement Models

Two basic kinds of errors affect empirical measurements: (a) random error, and (b) systematic error (Waltz et al., 1991). Random errors limit the degree of precision in estimating the true scores from observed scores that could lead to ambiguous measurement. Reliability is concerned with the extent to which measurements are repeatable. The amount of random error is inversely related to the degree of reliability of the measuring instrument (Carmines & Zeller, 1979).

The extent to which the measurement tool measures items other than the concept is referred to as systematic error. The validity of an instrument is a determination of the extent to which the instrument actually reflects the abstract concept being examined (Burns, 1987). As measurement of the concept improves, validity improves. As systematic error decreases, validity increases (Waltz et al., 1991).

There are two models used in the discussion of measurement error in classical measurement theory: (a) the domain-sampling model, and (b) the model of parallel tests (Nunnally, 1967). According to the domain-sampling model, any instrument is composed of a random sample of items drawn from a hypothetical domain of all items measuring the core concept (Nunnally, 1967). The score that any subject would obtain if it were possible to administer all items of the domain is the true score, and reliability of the measure is a reflection of the correlation of the sample of items with the true scores. If the sample of items correlates highly with the true score, the

reliability and validity of the measure will be high. The variability or consistency of items in sharing the core concept is reflected by the dispersion of correlations around the average (Nunnally, 1967).

Pedhazur and Schmelkin (1991) described reliability as the degree that test scores are free from errors of measurement. Using the domain-sampling model, reliability is based on the extent that the sample of items used correlates with the true score that would be obtained if all possible items in the domain could be measured. The measure of reliability is a reflection of the correlation of the sample of items with the true scores. If the sample of items correlates highly with the true score, the reliability of the measure will be high. If all of the items of the domain measure the core concept equivalently, the average of the correlations of each item with all the others (known as the reliability coefficient) is the same for all items. The variability or consistency of items in sharing the core concept is reflected by the dispersion of correlations around the average (Nunnally, 1967).

Norm-referenced measures are derived from classical measurement theory. In this view, the observed score is composed of a true score and an error score. Measurement error may be random or systematic (Waltz et al., 1991). For an instrument to measure the concept it purports to measure, it must be relatively free from error. Random errors of measurement affect reliability, while systematic errors decrease validity. When assessing the reliability and validity of the LTA-Care Acuity Index, the goal was to establish the degree of consistency in measurement as well as

the quality of the scores relative to what they profess to measure. Reliability and validity of empirical measurement are tools used to examine the concept of acuity in the long-term acute care hospitalized patient.

Assumptions

In conducting the study, the following assumptions were made:

1. The true score is that which would be obtained if there were no errors of measurement (Nunnally, 1967).
2. Obtained scores differ from true scores on a random basis because random measurement error is always present in the obtained score (Nunnally, 1967).
3. Random measurement error, a result of chance factors, is normally distributed. The mean of the error scores is zero, and the correlation between the true score and the error score is zero (Nunnally, 1967).
4. Increased measurement error results in an increased spread of obtained scores around the true score (Nunnally, 1967).
5. As random and systematic error decrease, reliability and validity increase (Burns, 1987).
6. Reliability and validity testing validates the use of an instrument for a specific group or purpose, rather than being directed toward the instrument itself (Burns, 1987).

Hypothesis

The research hypothesis was: The LTA-Care Acuity Index is a reliable and valid instrument for measuring acuity in an inpatient population: (a) the reliability coefficient alpha will be at least 0.70; (b) content validity will be measured by achieving a coefficient equal to or greater than 0.70 using the diagnostic content validation model; and construct validity will be established using known groups techniques and factor analysis.

Definition of Terms

The following terms were defined conceptually and operationally for this study:

1. Acuity was conceptually defined as a sampling of characteristics from a hypothetical domain of all items which measured the construct (Nunnally, 1967). For the purposes of this study, acuity was operationally defined as the cumulative score on the LTA-Care Acuity Index.
2. Classification system was conceptually defined as a norm-referenced measurement tool that reflected the amount of an attribute that a subject possessed. The operational definition for the purposes of this study was the LTA-Care Acuity Index.
3. Internal consistency reliability was defined as the extent to which all items on an instrument measured the same variables (Pedhazur & Schmelkin, 1991).

Operationally, internal consistency reliability was measured through the calculations of the coefficient alpha or Cronbach's alpha.

4. Long-term acute care was conceptually defined as a classification of hospitalization in which there was no set diagnostic-related group (DRG) restrictions imposed that limited the hospital stay; rather, length of stay was determined by criteria as set forth by Medicare. The operational definition of long-term care was the Medicare or Joint Commission on Accreditation of Hospital Organizations (JCAHO, 1998) designation of the facility as such.

5. Content validity was conceptually defined in terms of the domain sampling model, as reflecting the adequacy with which the domain was sampled. Operationally, content validity was defined as the mean correlation score calculated from submitted content validation forms completed by content experts using the diagnostic content validation (DCV) model.

6. Factor analysis was defined as the extent to which an instrument measured the intended concept through an evaluation of interrelated factors that underlie the instrument's set of items (LoBiondo-Wood & Haber, 1994). Operationally, factor analysis was reported as correlation coefficients that summarized the relationships between variables.

Limitations

The results of this study will be generalizable only to a population with characteristics similar to those characteristics of this study sample.

Summary

Health care costs continue to rise, and with the predicted nursing shortage within the next decade, it is ever more important to ensure the appropriate utilization of resources. While acuity tools have been developed for short-term acute care and the intensive care unit (ICU) patient population, there is little available in terms of any type of workload management systems in long-term care. Using classical measurement theory, this study examined the measurement of acuity in a long-term acute care hospital setting. The purpose of this study was to determine if the LTA-Care Acuity Index was a reliable and valid measure of acuity in this patient population.

CHAPTER 2

REVIEW OF LITERATURE

Employers and employees in the health care environment have had to adjust to major changes in their operating philosophies over the course of the last few years. Productivity and cost containment is the standard by which hospitals measure success. The surviving hospitals will be those that can best provide the necessary services in an effective and efficient manner to the health care consumer. A method of improving efficiency often involves changes in the number of full-time employees or FTEs. Nursing represents the major portion of the hospital's labor pool (Andro, Robertson, & Glandon, 1987), which was estimated at about 20% of both the hospital's operating budget and total cost of care (Dijkers & Paradise, 1986).

Hospitals are not only concerned about the cost for nursing services, they are more than ever before competing for patients and coping with managed care reimbursement caps. It has been reported that if some form of variable staffing is not used, a hospital will pay for 1 to 2 more hours of care per patient per day than is actually delivered. Therefore, an objective empirical basis for measuring productivity of a nursing staff becomes essential.

Workload Management Systems

The nursing workload is dependent upon patient care requirements and is directed toward the accomplishment of objectives established for that care (Aydelotte, 1973). A workload management system was a form of classification mechanism initially used in hospitals with few available resources for the development, implementation, and evaluation of a complex patient classification system. With a workload management system, direct and indirect nursing actions were quantified using several different approaches. Institutions generally used the approach that had traditionally been most successful in the recruitment and retention of caregivers. Carr-Hill and Jenkins-Clarke (1995) described the three major approaches to measuring nursing workload:

1. The dependency driven workload management system approach was based on the dependency needs of the patient and on a certain amount of nursing care required to perform the basic activities of daily living. Patient dependencies were combined with pre-determined timings of specific tasks, and unlike other approaches, differentiated between different types or levels of care.
2. The task-oriented approach was based on the recording and predicting of nursing interventions for individual patients. It was divided into two interacting components: activity and manpower. The activity component based costs on actual interventions, while manpower took into account salary structures that reflected experience and licensure. This approach was intended

to be a framework for monitoring planned and actual costs of all nursing services.

3. The care plan approach measured workload through the generation of nursing care plans derived from process and outcome standards for each patient. This approach did not quantify nursing actions in terms of the time required for task completion, nor did it attempt to correlate severity of illness with the amount of nursing time required for task completion.

The workload management system classified patients primarily according to their medical diagnoses. These somewhat general and unstructured measurement systems have advanced in form to the contemporary care maps and critical pathways that are frequently used in acute care hospital settings as well as the home health environment. While care maps and critical pathways are helpful in controlling costs, they tend to be focused on treatment of the disease as opposed to addressing all of the biopsychosocial needs of the patient.

Patient Classification Systems

As defined by Lewis and Carini (1984), a patient classification system--or patient acuity system--means the systematically identifying and assessing a group of patients' nursing care requirements on an individual basis. Giovanetti (1979) defined a patient classification system as a grouping of patients according to observable or inferred priorities or characteristics and nursing care requirements in order to

determine staffing needs. A patient classification system categorizes patients according to some assessment of their nursing care requirements over a specified period of time. These categories have included such terms as self-care, intermediate care, intensive care, long term or chronic care, home care and outpatient care. Over the last 30 years, patient classification has acquired a more specific meaning. It now includes categorizing patients according to an assessment of the acuity of illness, severity of symptoms, nursing dependency, and/or nursing interventions required (Alward, 1983).

When they were first developed and implemented in the hospital setting, the purpose of classification systems were to predict staffing needs and meet the regulatory requirements of agencies such as the Joint Commission on Accreditation of Hospital Organizations (JCAHO). What started out in the early 1980s as a recommendation by the JCAHO is now one of the 12 accreditation standards. Currently, all hospitals surveyed by this accrediting body must have a systematic method for planning and allocating human resources (staffing) that is based on a patient classification system (JCAHO, 1998).

Another incentive to establishing an effective patient classification system is that it is a method for determining nursing staffing patterns and the allocation of nursing personnel. As consumers became more conscious of health costs and as prospective reimbursement gained official support, efforts to cost out nursing services

for economic purposes have broadened the application of patient classification systems.

Patient classification systems were originally adapted from industrial time and motion models designed to quantify repetitive tasks, whose motions can be standardized, measured, and timed (Malloch & Conovaloff, 1999). The history of patient classification systems can be traced through at least three different evolutionary generations (Malloch & Conovaloff, 1999).

Before 1970, patient classification systems were calculated manually to estimate annual staffing needs and to project annual nursing budgetary estimations. Nurse-patient ratios were based on historical data. Staffing adequacy was monitored through quality and incident reports.

During the 1980s, hospitals began to feel the increasing presence of managed care. The federal implementation of diagnostic-related groups (DRGs) in 1983 dramatically affected the management of impatient services. Accrediting organizations mandated patient classification systems. Hospital administrators began looking at ways to control nursing costs while maintaining a quality standard of care as the cost of services increased and revenues decreased.

The 1990s saw increasing regulatory pressures continuing to challenge health care organizations. The Health Care Financing Administration's (HCFA's) 1997 revision in Medicare and Medicaid Conditions of Participation included a significant increase in staffing standards for operations. In addition, the JCAHO (1998) now

requires clinical outcome measures for accreditation. In several states, nurse staffing is a battle between a downsizing of the workforce and legislating mandated minimal staffing levels.

The present millennium will see a continued effort toward budget tightening and regulatory pressures. In addition, advances in technology should bring about automated medical record keeping and monitoring to the bedside.

Elements of a Classification System

Buckle, Horn, and Simpson (1991) described the essential elements of a viable patient classification system: (a) the system must be sensitive to multiple levels of variation in the physiological and behavioral status of the patient; (b) the system must be useful for all patients, regardless of location in the hospital; (c) the instrument should use objective criteria to measure variation in patient characteristics; and (d) the overall measure of patient complexity should be flexible and multi-purposeful. In addition, the system should be: (a) directly related to the time and effort spent on the associated activity; (b) economical and convenient to report and use; (c) mutually exclusive, so that no item is counted under more than one work unit; (d) open to audit, so that the accuracy of the work count is readily verified through the set up of a work count system or through existing internal work measurement programs or management information systems; (e) readily understood by those who plan, schedule, and control the work; (f) individually standardized as to procedures needed for

accomplishment; and (g) able to segregate registered nursing duties from other nurse staff requirements (Schroeder, Rhodes, & Shields, 1984).

Hospitals that use patient classification systems must have a method in place for validating the amount of care given to each patient for each unit. In addition, facilities must have a process in place for monitoring the reliability of the patient classification system over time (De Groot, 1989a).

Classification System Evaluation

The essential elements of a comprehensive classification system include: (a) simplicity, (b) efficiency, (c) objectivity, (d) acceptability, (e) utility, (f) reliability, and (g) validity. Each of these elements interacts with the four required nursing actions that drive successful classification implementation: (a) commitment, (b) coordination, (c) staff education, and (d) involvement. De Groot (1989b) used each of these components to develop a patient classification framework that allows for rapid problem identification and recommendations as well as assisting with the selection of an appropriate patient classification system (Figure 1).

The ideal patient classification system would match patient needs with the facility's nursing resources. In addition, the system would project short term as well as long term budgetary needs. Staffing needs would take into account the acuity of the patient, not just the skill mix of the nursing personnel. The patient classification

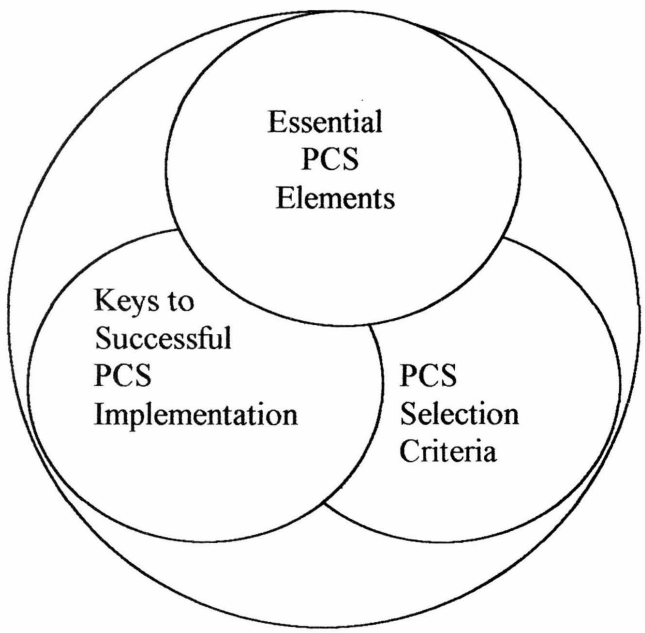

Figure 1. Theoretical Framework of a Patient Classification System (De Groot, 1989b)

system should be relatively simple and require a minimal amount of nursing time to classify patients (Alward, 1983).

Types of Classification Systems

The two most commonly used systems are the prototype evaluation and the factor evaluation (Reinert & Grant, 1981). In the prototype evaluation, categories are determined, parameters for each are defined, and the patient is assigned according to care needs. In factor evaluation, predetermined descriptors of care are defined, recorded separately for each patient, and then combined to determine the category of a particular patient. Descriptors of care may include activities associated with daily living, treatments, and psychosocial needs.

The prototype framework has three or four categories of patients, with descriptions of typical care requirements in each category. This type of tool is less commonly used because difficulties arise when a patient has care needs in different categories. The rater must establish which category will best fit the patient profile (Haas, 1988). This type of system may be difficult to monitor because it permits manipulation by the staff (Reinhart & Grant, 1981).

A factor evaluation instrument involves a list of critical indicators or descriptors of direct nursing care requirements which are checked separately; the sum of the indicators checked designates the patient's category (Haas, 1988). Factor evaluation appears to be more objective, but it may increase the paperwork and time involved for the nursing staff. Staff may not be receptive to becoming involved in patient classification and may question the meaningfulness of the evaluation (Reinert & Grant, 1981).

Two of the earliest known classification systems are the Commission for Administrative Services for Hospitals (CASH) and The Grace-Reynolds Application and Study (GRASP). The CASH system is an organization created by the Hospital Council of Southern California and Blue Cross to aid administrators in improving the cost effectiveness of hospital services through use of performance standards and system improvements (Georgette, 1970). In the early 1960s, a major effort of CASH was to conduct an in-depth survey of nursing service. From this survey, hospitals were selected to conduct studies measuring the total care and nursing service

requirements on various types of units. As a result, the workload system provided: (a) preplanning and scheduling of shift activities to prevent overloading or underloading one particular shift; (b) a patient centered care approach whereby all of the staff, regardless of position, participated in the rendering of care and services; and (c) the development of a patient care plan coupled with a personnel assignment schedule, which allowed nurses and nurses' aides to begin patient care without a lengthy verbal shift report. Because of this early survey work, the CASH patient classification system was developed. This system used a prototype design, and patients were classified into one of four categories according to type of illness, emotional status, medicines and treatments ordered, and general health (Des Ormeaux, 1977). Category I reflected patients requiring maximum care; Category II reflected mostly medical and surgical patients; Category III reflected patients needing less than average care, and Category IV reflected patients requiring minimal care (Schroeder et al., 1984). The number of care hours needed were then calculated from the number of patients in each category. This system was widely adopted and used by several by California hospitals.

The Grace Hospital in Morgantown, North Carolina developed the Grace-Reynolds Application and Study (GRASP) patient classification system in the early 1970s. The scores on the GRASP instrument are associated with standard care times for specific nursing interventions. The number and scope of indicators vary by hospital and average more than 45 items. The tool calculates the patient's total

nursing care need by summing the care activities listed on the instrument and adding a fixed factor for indirect care (Phillips, Castorr, Prescott, & Soeken, 1992).

Hospitals that choose to use factor evaluation of the GRASP system may either choose to conduct their own internal time studies or use the item values provided by the GRASP system consultants. Total care minutes are converted into patient care units that represent one hour of required care. Because both the nursing interventions and associated time standards are specific to a given institution, each institution must conduct its own reliability and validity testing (Phillips et al., 1992).

Current Trends in Patient Classification

In the process of evolving patient classification systems, several notable changes have surfaced that have somewhat standardized the format of patient classification from qualitative to quantitative. Indexes based on physiologic data have been intentional efforts designed to upgrade the reliability, validity, and standardization of the measures (Aronow, 1988; Iezzoni, 1990; Wagner & Draper, 1984; Wagner, Knaus, & Draper, 1989). Second, emphasis on expansion of scope beyond the medical model has included measures of functional status, measures of independence in self-care, and measures of psychosocial function which have become recognized patient care needs.

Limitations of many of the existing nursing classifications systems are widely recognized by nurses and the developers of the various systems (Haas, 1988).

Limitations include the following (Buckle et al., 1991): (a) existing systems do not capture patient-specific characteristics; (b) most systems are designed with a task-orientated approach as opposed to being a tool to objectively evaluate patient needs; (c) due to variations in clinical privileges from state to state and institution to institution, no one nursing classification system has been widely accepted and tested; (d) the content of many classification systems is subjective and not sufficiently sensitive to important patient characteristics; (e) a system has not been developed that identifies costs associated with the delivery of nursing care; and (f) universally accepted standards of nursing care have not been established. For these reasons, a generally accepted classification system for nursing is not feasible, and many institutions are developing their own systems to predict patient acuity and staffing needs.

The development of internally based classification systems has its advantages, but it is not without major potential weaknesses. In 1988, Nagaprasanna surveyed 251 for profit and not-for-profit hospitals. Only 16% of the hospitals surveyed were using a commercialized patient classification system. Forty-two percent of the facilities reported using some form of internally produced instrument. While ease of classification was the most important factor in the selection of an existing system, system acceptance was the greatest contributing factor in internally created systems (Nagaprasanna, 1988).

Measurement of the reliability and validity of the instrument is especially problematic for the hospitals that adopt existing systems with little or no modification in the categorization or quantification schemes for the particular nursing service (Alward, 1983). Variables can affect the instrument reliability and validity from one setting to the next. Because the establishment of the reliability and validity of the patient classification system must take into account the institutional philosophies, care delivery systems, skill mix, supplies, medical staff expectations, and the availability of support services, the validation results may be considered specific to individual institutions (Alward, 1983; Giovanetti, 1979).

Research Related to Patient Classification Systems

The exact number of patient classification systems in existence is not known. Malloch and Conovaloff (1999) reported some 40 types of patient classification systems. An unofficial count over a decade ago revealed some 1,000 hospitals using some form of patient classification system (Nagaprasanna, 1988). Each varied in format and approach to classification.

The Rush-Medicus patient classification tool consisted of 37 critical indicators reflecting the care required by each patient. These indicators were given a weighted point count and total points were summed for each patient. Some indicators were very specific and objective (such as invasive monitoring) while others were more subjective, such as sensory defects (Phillips et al., 1992). Medicus scores were based

on summation of pre-determined weights for each indicator. Weights were derived by direct observation of nursing activity, work sampling, and consultation (Medicus Systems, 1983). Total scores were translated into one of five acuity levels that ranged from Type I patients who required minimal supportive nursing care to Type V patients who required the equivalent of one-on-one nursing. Phillips and co-workers noted that although it was reported that years of reliability testing had been undertaken, no specific results were addressed in the literature. However, Batty, Mooney, and Lowry (1990) reported interrater reliability coefficients of greater than 0.90.

A critical care patient classification tool developed by the University of California San Diego Medical Center used the factor evaluation method containing varying numbers of critical indicators based upon the type of unit for which the tool was designed (Niemeier & Reed, 1985). Patients were prospectively classified each shift by the staff nurse caring for the patient. A summary of the numbers of patients in each classification for each nursing unit was then used by the nursing staffing office to determine and allocate nursing personnel. One of the major drawbacks to this system was the time involved in patient assessment. It took about one hour to assess three to five patients. This amount of nonproductive patient care time could increase nursing costs as well as add to overall job dissatisfaction, as staff could potentially fall behind schedule in their patient care.

Interrater reliability was done on the Medicus system monthly on alternating shifts. If it became apparent that downward trends in reliability coefficients occurred,

interrater reliability was measured and reported weekly. Because interrater reliability can be subjectively affected by the attitude of the staff completing the tool, the results may have actually been more influenced by the time demands imposed by the instrument itself. Although estimates of validity commonly accompany reliability statistics, there was no indication that they were performed on the instrument.

One of the more recent classification systems that is undergoing pilot testing in various acute care settings is termed the 3PCS (Malloch & Conovaloff, 1999). This system includes five elements: (a) standardized, research-based interventions and outcomes categories; (b) descriptions of patient care in comprehensive units of service; (c) identified caregiver roles; (d) patient medical record documentation forms; and (e) caregiver competency profiles. It is unclear how each of these elements is measured or tested in terms of their psychometric properties. The evaluation process included a review of charting compliance and interrater reliability as the methods compared to the data contained on a productivity management spreadsheet. The authors stated that the system was piloted in an acute care setting once the reliability and validity targets were met; however, the methods of validity testing were not explained. None of the staff in any of the pilot settings had formally evaluated the 3PCS. Instead, the intent was to evaluate the tool through input from staff members, administration, patients, and families.

The Acute Physiology and Chronic Health Evaluation Tool (APACHE) (Wagner, Knaus, & Draper, 1983) was one of the first medically driven acuity-based

classification systems. The instrument was primarily designed to assess the severity of illness and probability of mortality for patients admitted to ICU. The original system was based upon 34 physiological variables, reflecting the degree of insult to the seven vital systems of the body as well as evaluating the presence of chronic health problems. The APACHE system assigns a numeric score ranging from 0-71; the greater the score, the more severely compromised the patient. Second and third generations acuity scoring systems, APACHE II and APACHE III used fewer physiological variables but added the additional variables of age and chronic health evaluation to predict patient outcomes (Wagner & Draper, 1984; Zimmerman et al., 1998).

Patterned after the medical model, one of the most widely used systems developed to classify patients is the Therapeutic Intervention Scoring System (TISS). The TISS was devised at the Massachusetts General Hospital (MGH) to provide quantitative data to justify nursing staffing in intensive care units. It has been used as a workload management system as well as for cost accounting in the ICU setting (Jackson & Resnick 1982). This method attempted to classify severity of illness by quantifying therapeutic interventions. A committee of intensive care physicians and nurses assigned point values according to the time and effort required for nursing care (Cullen, Civetta, Briggs, & Ferara, 1974). Massachusetts General Hospital used the scores from original TISS instrument to place patients in one of four classifications based upon the number of points accrued. Class I patients averaged 5 ± 0.2 points;

Class II patients averaged 11 $\pm$ 0.7 points. Class III patients averaged 23 $\pm$ 1 points; and Class IV patients averaged 43 $\pm$ 1 points. Nursing hours per patient day were then determined by the total number of points/patient. For example, a patient with a score of 43 would need one-on-one nursing, while an intermediate care patient averaging 12-13 points would require a nurse/patient ratio of 1 to 4 (Jackson & Resnick, 1982).

The origina 76-interventions TISS form was subjected to a major revision in 1983 (Keene & Cullen, 1983). Items were deleted or added, and certain items had a point score adjustment. The revised form was then evaluated on 100 patients in three separate ICUs. A regression equation demonstrated that the 1983 system was equivalent to the 1974 system. Additional psychometric testing, if done, was not reported (Wagner et al., 1983).

In 1996, Moreno and Morais (1997) modified the 76-item TISS and developed the Simplified Therapeutic Intervention Scoring System (TISS-28). Reliability and validity studies compared the TISS 76 to the simplified version. The results indicated that the TISS 28 was a reliable and valid instrument for the measurement of nursing workload. The study was limited in scope as the instrument was tested in Portuguese intensive care units.

Clermont, Angus, Linde-Zwirble, Lave, and Pinsky (1998) generated a computerized version of the TISS as an index of resource that could be used in various types of ICU settings. The researchers developed an automated mapping of

the hospital billing database into the different items of the TISS and generated computerized active TISS scores on 1,372 ICU days. Trained data collectors then validated the computerized score by comparing it to prospectively gathered active TISS scores. There was a significant positive correlation between the scores on the TISS from the billing data and the computer-generated TISS scores, thus placing a correlation between the utilization of hospital resources and severity of illness.

The format of the TISS was used to develop a patient classification system for use in a South African ICU patient population. The use of the original TISS outside of the United States was invalid because of differences in critical care units. As a result of shortages, the nursing responsibilities of South African registered nurses (RNs) differ from their American counterparts. In addition, support staff such as respiratory therapy, occupational and physical therapy, and social workers either do not exist or have limited responsibilities in terms of the overall care of the patient. A shortage of medical personnel in South African critical care units places extra responsibility on the nurse to perform physician-orientated procedures.

The CritScore (Scribante, Muller, & Lipman, 1996) was patterned after the TISS in terms of the scoring system. Items not appropriate in this patient population were eliminated, while items such as nursing care of the disoriented patient and counseling of patients and/or families were added to the instrument. Reliability and content validity were done on the CritScore; however, the statistical results were not reported.

The LTA-Care Acuity Index

The strength of internally developed classification systems is in the relationship between the variables contained within the instrument and the characteristics of the patient population for which the instrument is used. An instrument that is not reflective of all of the patients' nursing care needs will not be a reliable and valid measurement tool. Most of the commercially produced patient classification systems are used in a short-term acute care setting and are focused primarily on observable and quantifiable nursing care activities. These systems do not take into account many of the psychosocial implications of a long-term hospital stay on both patients and families. In addition, none of the instruments reviewed took into account the multiple medication administrations that are given to patients with multiple system illnesses.

The LTA-Care Acuity Index contains 77 variables divided into 11 subscales. Six of the subscales are body systems and include pulmonary, cardiovascular, gastrointestinal, renal/genitourinary, integumentary, and neurological. Other subscales encompass mediation administration, intravenous therapy, daily activities, special procedures, and psychosocial variables. In addition, medication points are factored into the total acuity score by adding the total number of medications that the patient received in a 24 hour period to the total number of activity points. This number is the total acuity score for that patient.

The design of the LTA-Care Acuity Index instrument contains the following features:

1. It is a point system based upon patient needs rather than diagnosis or subjective assessment of patient acuity.
2. It is adaptable to future changes and computerization and capable of assuring a high degree of reliability that could be monitored for accuracy.
3. Nursing management can use the system for staffing and budgetary purposes, as it can be accurately correlated with nursing time required for specific patient care activities and compatible with established staffing patterns throughout the hospital.
4. Completion of the instrument takes a minimum amount of time and effort.
5. The system lends itself to spot checking and monitoring as well as modification when indicated by future changes in acuity of the long-term acute care population.
6. The tool that allows for multiple administrations enables the nursing administrator to justify changes in nursing care needs for each of the patients.

Regardless of the type of instrument, an accurate classification system predicts the number of personnel required to meet patient care needs. The system should enable the nurse to qualify and quantify individual patient needs so that adequate and safe levels of nursing care can be provided. When patients always receive exactly the care they need at the appropriate time, health care will have achieved its ultimate altruistic goal: a holistic, humanistic, and seamless integrated health care delivery system (Malloch & Conovaloff, 1999).

No one system can be judged a superior acuity tool unless effective quality control procedures are put into place to ensure accuracy of measurement. Errors in classification can occur in several ways. Misclassifications may be due to improper use of the instrument, a nurse's classification of patients without sufficient knowledge of their needs and conditions, or inflation of patients' classification levels simply to get more help (Niemeier & Reed, 1985). Initial and regularly scheduled reliability and validity checks are essential quality control tools in the implementation and evaluation of a patient classification system.

Psychometric Testing

Giovanetti and Mayer (1984) purported that establishing and maintaining system reliability and validity are vital to acceptance of a system and the ability to use the information with confidence. It must be known whether or not a particular patient classification tool actually predicts different amounts of care for individual patients or groups of patients. This information is vital in assuring the reliability and validity of the classification tool in a given setting (De Groot, 1989b). Even if an instrument has been shown to be reliable and valid in one setting, this does not ensure that, in a different environment, the tool will maintain its psychometric integrity. The amount of nursing care received by the different levels of patients can vary considerably from institution to institution and between units within a single facility. This variation is due to such factors as patient characteristics, existing staffing levels, physical plant

consideration, types of ancillary services, and levels of staff (De Groot, 1989b). Yet, vigorous testing of the validity and reliability of instruments that measure nursing care has frequently been found lacking (Whitney & Killien, 1987).

Measurement is the process of translating reality into numbers (Knapp, 1985). Reliability issues arise when the fit between the true score and the obtained score is studied, whereas validity issues arise when the fit between the construct and the true score is studied (Knapp, 1985). There are three primary concepts involved in translation process: (a) the construct (C), (b) the true score on the variable (T), and (c) the obtained score on that variable (X) (Figure 2).

Construct (validity)	True Score (reliability)	Obtained Score
C	T	X
Acuity	The mean score on the LTA-Care Acuity Index for all Long-Term Acute Care Hospitalized Patients	A score on the LTA-Care Acuity Index

Figure 2. A Conceptualization of Reliability and Validity of the LTA-Care Acuity Index (adapted from Knapp, 1985).

Reliability

Reliability is concerned with how consistently the measurement technique measures the concept of interest (Burns & Grove, 1993). Reliability is not a stable property; it not only can, but usually does, change with each use of an instrument.

Therefore, estimates of reliability are specific to the sample being tested (Lynn, 1989). Reliability testing needs to be performed on each instrument used in each study sample prior to performing other statistical analysis, as estimates of reliability reflect the accuracy of the data collected.

Reliability is often viewed as the accuracy of an instrument and is the ratio of desired information to the obtained information or the extent to which measurement error is minimized (Nunnally, 1978). Theoretically, reliability coefficients range from 0.00, or no reliability, to 1.00, or perfect reliability. Because errors of measurement always occur in research, reliability is consistently less than the ideal value of 1.00 (free of measurement error) (Lynn, 1989). A reliability coefficient of 0.70 is considered acceptable for a newly developed instrument, and a reliability coefficient of 0.80 is the lowest acceptable value for an established instrument (Nunnally, 1978).

The reliability of any instrument is best established by measurement of three components: (a) equivalence--the extent to which instruments measure the same traits in the same subjects; (b) stability--the extent to which the same results are obtained on repeated administration, and (c) internal consistency or homogeneity--the degree to which all subparts measure the same characteristic (Ebner, 1985).

The aspect of equivalence is focused on comparing two versions of the same instrument or having two observers measure the same event. Statistical correlation techniques are used to provide an estimate equivalence reliability; the minimum acceptable value is 0.80 (Burns & Grove, 1993).

The measurement of the stability of the instrument is desired when an instrument is reviewed for its appropriateness in the setting in which it is being used. Assessments of instrument stability are determined through test-retest reliability procedures that yield a reliability coefficient.

Internal consistency is one of the most frequently generated estimates of reliability for instruments composed of a number of items or variables that will be formed into a linear composite (Ferketich, 1991). Internal consistency is used to determine the extent to which each item in the instrument measures the concept. Measures of internal consistency reliability are more frequently used and subsequently reported in the literature for several reasons: (a) the subject is not burdened with the completion of an alternative form of the instrument, (b) the subject does not need to retake an instrument a second time with all the attendance problems, and (c) the researcher is not required to deal with the arbitrariness of split-half procedures. It is expected that each item reflects the concept to be measured, and as such, all items are highly correlated.

Internal consistency reliability is measured with Cronbach's coefficient alpha. The alpha estimate uses all of the information about the variance and covariance of the items and therefore has many desirable properties for estimating the reliability of the multiple item instrument (Ferketich 1991). Cronbach's alpha coefficients range from 0.00 to 1.00 indicating a low to very high internal consistency. The strength of inter-item correlation is reflected in the alpha score. Many researchers consider an alpha

coefficient of at least 0.70 to be adequate for an instrument in early stages of development and a coefficient of at least 0.80 to be adequate for a more developed instrument (Nunnally, 1978; Polit & Hungler, 1987). Very high alpha coefficients are difficult to obtain in test development and may be indicative of redundancy among items (Nunnally, 1978). Redundancy can be assessed though the examination of the correlation matrix. When inter-item correlations are consistently above 0.70, redundancy may be a problem. Alternatively, when inter-item correlations are consistently below 0.30, there may be a lack of substantive relation among the items measuring the construct (Nunnally, 1978).

Validity

Validity of an instrument is established when the instrument (representing an operational definition of a property) actually measures the conceptually defined property it is intended to measure (Polit & Hungler, 1987). The validity of an instrument is a more important consideration than the reliability once the variable of interest has been defined and operationalized. Unlike reliability, validity can be a stable property of an instrument as long as the instrument is used in the same manner for which it was developed (Lynn, 1989). When planning to validate an instrument, a systematic process for data gathering and analysis is required to ensure that validation evidence is strong (Slocumb & Cole, 1991).

Validity is a judgment about the meaning of what is estimated by an instrument and serves as a guide for determining the appropriate implication of findings. Validity data are referred to as evidence rather than proof, because uncertainties exist even in well-developed, broadly field-tested instruments (Slocumb & Cole, 1991). The three primary types of validity are: (a) construct validity, (b) criterion-related validity, and (c) content validity.

Construct validity concerns the appropriateness and adequacy of those variables within the tool as well as how they are operationalized and quantified (Shelly, 1984). The variables or critical indicators that establish how patients are classified in acuity instruments are all constructs developed from theory and empirical data (Ebner, 1985). The types of nursing actions, medications, patient assessments, educational activities, and self-care abilities are all examples of variables contained in a patient classification system. In measuring construct validity, it is necessary to verify these constructs thorough successive modification of the instrument, until the tool eventually measures what it is supposed to measure. Few attempts toward establishing construct validation for any patient classification system have been made, primarily because of the confusion about what these instruments actually measure (Alward, 1983).

Criterion-related validation does not address how well an instrument is measuring a particular trait; rather, it is a predictor of a relationship between the instrument and some criterion. Criterion validity is of two types: (a) concurrent and (b) predictive. Concurrent validity is assessed by comparing results for the new

instrument with an existing instrument that has established validity and reliability. The difficulty in using this type of validation technique is related to the lack of established instruments that measure acuity in the long-term acute care hospital setting. Predictive validity measures the extent to which predicted nursing care requirements reflect the actual care delivered. To predict future performance of the instrument, it is first necessary to establish the concurrent and content validity of the instrument.

Content validity refers to the sampling adequacy of the content area or domain being measured (Polit & Hungler, 1987). Content validity involves the use of individuals considered knowledgeable about the content area. Construct and criterion validity data are gathered from a sample of respondents for which the instrument was designed; subsequent analyses are most often statistical. In contrast, analysis of content validity is focused on an assessment of judgments provided by experts (Lynn, 1986).

The judgment of experts, that is, experienced staff nurses, is used for determination of content validity (Williams, 1988). Patient classification systems' content validity is usually assessed by having nurse administrators and clinicians from several different clinical areas review the tool. Nurse administrators with experience in using patient classification systems may also be used as content experts to offer further general support for the instrument's content validity (Ebner, 1985). Analysis of content validity is focused on an assessment of judgments provided by experts. It requires that experts confirm that the operational definitions created are congruent

with the universe of the theoretical definition of the concept (Slocumb & Cole, 1991). Expert judges should be consulted after the items (operational definitions) have been developed. Input from staff nurses who will be using the tool is critical (Ebner, 1985).

Determining the number of experts needed is dependent upon the number of experts that the researcher can identify, not on a population estimation principle (Lynn, 1986). Lynn detailed specific guidelines that can be applied to the selection of experts for content validity determination (Table 1).

Table 1

Proportion of Experts (Above the Line) Whose Endorsement Is Required to Establish Content Validity Beyond the 0.05 Level of Significance (Lynn, 1986)

Number of Experts	2	3	4	5	6	7	8	9	10
2	1.00								
3	0.67	1.00							
4	0.50	0.75	1.00						
5	0.04	0.06	0.8	1.00					
6	0.33	0.50	0.67	0.83	1.00				
7	0.29	0.43	0.57	0.71	0.86	1.00			
8	0.25	0.38	0.50	0.63	0.75	0.88	1.00		
9	0.22	0.33	0.44	0.56	0.67	0.78	0.89	1.00	
10	0.02	0.30	0.4	0.05	0.60	0.70	0.80	0.90	1.00

Although the number of experts that should be included in the validation of any instrument is determined by the researcher, it is recommended that during the development phase, a minimum of 5 domain experts be used to control for chance agreement, and a maximum of 10 experts be used for validation of the instrument. The use of only two judges is not only statistically unjustifiable, it can place the instrument developer at great risk of erroneous conclusion that content validity has not been achieved when it actually has (Lynn, 1986).

One type of quantification method for the determination of content validity is the Content Validity Index (CVI) (Waltz & Bausell, 1981). The CVI is formatted in a 4-point ordinal rating scale. The calculated CVI score is applied to each of the instrument's individual items as well as a score for the entire instrument.

The Diagnostic Content Validation Model developed by Fehring (1987) uses a similar Likert-type formatting for quantification of content validity. It differs from the CVI in that it uses a 5-point scale with responses ranging from "not very characteristic" to "very characteristic" for each of the items being evaluated. Weighted ratios are applied for each of the responses to each of the items included on the scale. Weights are provided so that the total score can reach a maximum of 1.00. This type of scoring prevents a value being given to a defining characteristic that the content experts judge not to be representative of the construct being evaluated. The ratio scores are summarized and divided by the number of content validity respondents. The result is a calculated correlation coefficient for each of the items

contained in the instrument as well as an overall DCV score. Furthermore, defining characteristics can be individually examined for their weighted ratios and labeled as major or minor characteristics. A major defining characteristic is one in which the overall content validity coefficient is 0.80 or greater (Polit & Hungler, 1987).

Factor Analysis

Factor analysis is used to simplify complex sets of data (Kline 1994). A factor is defined as a dimension or construct which is a condensed statement of the relationships between a set of variables (Kline, 1994). The basic assumption of factor analysis is that underlying dimensions, or factors, can be used to explain complex phenomena. Observed correlations between variables result from their sharing these factors.

The type of factor analysis is dependent upon the measurement model. Classical measurement theory details that all measurement error is random, and therefore, all variance is unique to an individual item and not shared with any other item or factor in the instrument (Nunnally, 1978). Principal components factor analysis, consistent with classical measurement theory, assumes that all error is random and not reflective of an underlying structure (Ferketich & Muller, 1990). If all error is random and the average of the error sums is zero, each variable (item) would correlate perfectly with itself and would have a correlation coefficient of 1.00 (Ferketich & Muller, 1990).

Factor analysis usually proceeds in four steps (Kim & Muller, 1978; Kline, 1994; Norusis, 1990):

1. The correlation matrix for all variables is computed. This is generally termed factor loading. Variables that do not appear to be related to other variables can be identified from the matrix and associated statistics. A factor loading of 0.30 indicates that 9% of the variance is accounted for by the factor and is generally considered significant.
2. The second step of the process is termed factor extraction; the number of factors necessary to represent the data. The goal of factor extraction is to determine the factors. The method of principal components analysis creates linear combinations of the observed variables. The first principle component is the combination that accounts for the largest amount of variance in the sample. The second principal component accounts for the next largest amount of variance and is uncorrelated with the first. Successive components explain progressively smaller portions of the total sample variance, and all are uncorrelated with each other. This total amount of variance is termed the eigenvalue for the factor.
3. The third step of factor analysis is termed the rotation phase. Most factors are correlated with many variables. The purpose of the rotation phase of factor analysis is to transform the initial matrix into a more simplified structure. The varamax orthogonal rotation attempts to minimize the number of variables that

have high loadings on a factor, with the intent to enhance the interpretability of the factors.

4. In step four of factor analysis, scores for each factor can be computed for each case. Factor scores are estimates for a case on an underlying factor formed from a linear combination of observed variables. Factor scores can be used to represent the values of the factors.

Instrumentation Reports in the Literature

Lynn (1989) reviewed 20 research articles published in issues of Heart and Lung, the Journal of the American Association of Critical Care Nurses. Of the 20 articles, 17 questionnaires and 17 different physiologic or technologic measures were used in data collection, for a total of 34 methods of instrumentation. For the studies in which questionnaires were used, only 2 of the 17 reported any psychometric testing of the instrument before, during, or at the conclusion of the study. Some researchers reported reliability and validity coefficients that had been obtained with the use of the instrument in previous studies at other research sites. Other researchers reported the reliability and validity results based upon the developer's own claims. According to Lynn, the quality of the instrument cannot be determined unless at least minimum descriptive, reliability, and validity information is available to support it. If the instrument is not tested, the data and the results should not be assumed to have merit.

Summary

Patient classification is the process of categorizing patients according to an assessment of their individual care requirements. The assessment should quantify the patient's biopsychosocial and biophysiological needs as well as patient education and nursing procedures. The tool must be adaptable to most clinical areas and must be accepted by the hospital administration as well as the nursing staff. Moreover, the patient classification tool must be valid and reliable, while accurately predicting the nursing workload.

Historically, patient classification systems have been utilized most significantly to organize nursing staff in hospital settings. Information that can be provided by a reliable and valid system can also help determine appropriate staffing mix, plan daily nursing care assignments, prioritize patient care needs, and place patients where the most appropriate nursing and material resources are available. A classification system can also justify overtime hours and decisions to either fill vacant positions through active recruitment programs or to leave positions open temporarily. The information also can help monitor patient care to assure its quality and assist with the discharge planning process. Acuity scores provide a justification system for adding prn or agency nurses and lend support to the need for hospitals to float nurses to other units. Daily patient assignments based upon a classification system can be more equable. Finally, a patient classification system can produce information to guide budgetary planning and allocation and serve as a basis for variable billing systems (Reitz, 1985).

Haas (1988) recommended that certain issues be addressed with regards to patient classification instruments so that time demands for each patient can be predicted over an entire hospital stay at the outset: (a) vigorous ongoing checks on reliability, (b) determination of validity of the instrument, (c) expansion of the factors which delimits the model for each category, (d) increased sensitivity to patient differences, and (e) enhanced predictive ability. From the review of literature, it appears that these issues are the forefront of the development, implementation, and evaluation of a classification system.

CHAPTER 3

PROCEDURE FOR COLLECTION AND TREATMENT OF DATA

An exploratory, nonexperimental, methodological study was undertaken to describe the development validation of a scientific patient classification instrument (LTA-Care Acuity Index) for long-term acute care patients. A descriptive research design was appropriate as there was a lack of information with respect to the variables of interest in a given population and because these variables were not amenable to manipulation (Brink & Wood, 1989). An instrumental research design involving two phases was used for LTA-Care Acuity Index development. The development phase consisted of two steps: (a) domain identification, and (b) item generation. The literature review identified the scope and nature of the concept of acuity (domain identification), and a provisional instrument was developed (item generation).

The second phase of this study consisted of the judgment or quantification phase. Inclusive of this phase was the reliability and validity testing of the instrument.

Setting

Data were collected from a long-term acute care hospital in a large metropolitan area. The facility is licensed for 110 beds with 5% of the total number of beds designated as ICU. Ten percent of the beds in this facility comprise an

intermediate care unit (IMU). The remaining 85% of the hospital beds are licensed as general medical-surgical beds. The primary admitting diagnoses are: respiratory failure, dependence on mechanical ventilation, and functional decline. The average length of stay is 42 days.

Population and Sample

The accessible population was all patients admitted to a long-term, acute care facility. The sampling frame consisted of all patients admitted to the long-term acute care hospital between 1/1/00 and 6/1/00. There were no restrictions regarding patient inclusion, as every patient received initial and weekly acuity scores as a function of routine nursing staff responsibilities.

A nonprobability consecutive sampling technique was used for data collection. Random selection was not feasible for this study due to the large sample base required for instrument testing. Despite the disadvantages of nonprobability sampling, including the risk of bias, this type of sampling design has reported advantages of being practical and economical (Pedhazur & Schmelkin, 1991). Consecutive sampling was appropriate, because there were no exclusion criteria.

Sample Size

Power analysis was used to determine the needed sample size for hypothesis testing. The statistical power of a test is a function of four factors: (a) the significance criterion (a), (b) the variance of the data (s^2), (c) sample size ($\underline{n}$), and (d)

a factor that reflects the magnitude of the observed differences (effect size or ES) (Portney & Watkins, 2000). Power analysis for correlations is based on the magnitude of association, or the correlation coefficient (r) with the effect size index being the value of r (Portney & Watkins, 2000). Cohen (1988) addressed the r as it relates to the correlation coefficient when no other statistical rationale is available: small r = 0.10, medium r = 0.30, large r = 0.50. The minimum sample size needed for reliability testing generating a correlation coefficient, r (a_2 = .05) with a power of .80 and an ES of .30 was 153 (Cohen, 1988).

Protection of Human Subjects

Guidelines of the Human Subjects Review Committee at Texas Woman's University-Houston Center were followed to assure protection of the study participants. Approval was obtained from the University as well as the institution at which the research was conducted (Appendix B).

This study used an observer-rated tool. Because the research involved data routinely collected by nursing staff members, written informed consent was not obtained from the patients. Acuity tools are not part of the patient's permanent medical record, and LTA-Care Acuity Index scores are not used as a provision of patient care.

Patient confidentially with the potential for disclosure of the names and their identifying clinical information was a potential risk to the subjects involved in this

research. To ensure patient confidentiality, only copies of acuity tools were provided to the researcher and entered into the database for statistical analysis. All patient identifiers with the exception of patient location (2nd, 3rd, or 4th floor) were blacked out. This process eliminated any mechanism for associating an acuity score with individual patient data.

Reporting all results in aggregate form that in no way identified individual subjects protected the rights of the study subjects. All research materials were kept confidential and only accessible to the researcher.

Instrument Testing

Within classical psychometrics, two of the most important aspects and prerequisites of a test are its reliability and its validity (Rust & Golombok, 1989). Pedhazur and Schmelkin (1991) described reliability as the extent to which a measurement is consistent and free from error. Reliability can be conceptualized as reproducibility or dependability (Portney & Watkins, 2000). The second prerequisite, validity, ensures that a test is measuring what it is intended to measure. Validity is necessary for drawing inference from data and determining how the results of an instrument can be used (Portney & Watkins, 2000). Validity implies that a measurement is relatively free from error; that is, a valid test is also reliable.

Internal Consistency Reliability

Reliability, reported as correlation coefficients, is the amount of stability, internal consistency, or equivalence of a measurement tool. Internal consistency or homogeneity refers to the extent to which different parts of an instrument are equivalent in measuring the attribute being studied. Procedures to evaluate internal consistency are economical because they require only one test administration. They are also the best method to assess one of the most important sources of measurement error, the sampling of items (Polit & Hungler, 1987; Shelley, 1984). The method chosen for evaluating internal consistency was the coefficient alpha or Cronbach's alpha.

Cronbach's alpha coefficient is the most commonly used and preferred index of internal consistency reliability (Waltz, Strickland, & Lenz, 1991). The alpha coefficient measures the degree to which responses on any item correlate with responses on any other item in the same instrument at a single administration of a test. Cronbach's alpha takes into account both the average correlation among items and the number of items (Nunnally, 1978).

Although determination of a tool's reliability is necessary, it is seldom sufficient for evaluating the quality of scores (Summers, 1993). Validity of an instrument reflects to what degree the tool actually estimates the type of characteristics it is supposed to estimate (Whitney, & Killien, 1987). When applied to the LTA-Care Acuity Index, the question becomes, "Does the tool accurately measure the nursing

care requirements of patients?". There are several types of validity that can be used in the development of an acuity tool. Instrument validity was determined for this study through content and construct validity.

Content Validity

The methodology for obtaining content validity through an expert panel used the criteria described by Fehring (1987). The expertise of the panel was determined through educational background as well as clinical experience (Slocumb & Cole, 1991). The purpose of content validation is to have content judges offer expert opinions that support or reject the adequacy of each of the items contained in the construct. A minimum of 10 experts including nurses in administrative as well as clinical roles were asked to review the acuity tool. Generally, only nursing experts are chosen because of their knowledge about categorizing nursing needs. Ebner (1985) stated that input from staff nurses who were using the tool was critical.

A personalized letter was sent to the prospective members of the expert panel explaining the purpose of the instrument and instructions for determining the content validity of each of the items in each of the instrument's 11 categories. A key, with an explanation of the point value system was included with the instructions for completion (Appendix C). Additional information which was presented for evaluation of content validity by the members of the expert panel included a content validation

form, a summary of the pertinent literature with respect to patient acuity instruments, and instructions for rating the items in order to define the concept.

Content experts had 30 days to complete the evaluation of the instrument. A self-addressed stamped envelope was included to facilitate return.

Construct Validity

Consistent with classical measurement theory, the principle component factor analysis model contends that all measurement error is random, and therefore, all variance is unique to an individual item and not shared with any other item or factor in the instrument (Ferketich & Muller, 1990). In determining construct validity, factor analysis examines the structure within a large number of variables and attempts to explain the nature of their interrelationships (Kim & Mueller, 1978). Factor analysis examines the structure within a large number of variables in an attempt to explain the nature of their interrelationships (Kim & Mueller, 1978). Factor analysis is classified as exploratory or confirmatory (Nunnally, 1978). Exploratory factor analysis is used in initial instrument development to determine which items best represent the concept under study. This procedure creates a correlation matrix for all the test items. Based on these correlations, the factor analysis attempts to identify the principal components of the data through factor matrix. The factor matrix contains the factor loadings for each variable on each factor. Factor loadings greater than 0.30 are generally considered indicative of some degree of relationship. For the purpose of

this research project, factor analysis attempted to identify the principal components of the data and categorize sets of variables that were linearly correlated with each other.

A one-way analysis of variance (ANOVA) was calculated to identify differences in characteristics of known groups. Significant known-group differences were further identified through a Tukey's post hoc test.

Data Collection

After agency and institutional review board approval was obtained, data collection began. All patients admitted to the facility have an acuity score calculated within 24 hours of admission. Nursing staff complete subsequent patient acuities on a weekly basis. Exceptions requiring additional acuity scoring included a change in a patient's condition that resulted in a move of that patient to another level of care. In each case, the second administration of the acuity tool was chosen for reliability and validity testing, as the initial administration was done upon admission and was generally not reflective of daily, routine nursing care.

Treatment of Data

Data were coded to ensure confidentiality by using a table of random numbers. Demographic data were collected to categorize subjects by level of care (1 = ICU, 2 = IMU, 3 = 3rd floor general medical-surgical unit, and 4 = 4th floor general medical-surgical unit). Data were entered into a computer and checked for accuracy.

All descriptive and inferential statistics were calculated using SPSS® Version 9.0 for Windows© statistical software package. Descriptive statistics included frequency distributions, measures of central tendency, and variability. Inferential statistics were used to compare means and address the hypothesis as to whether the LTA-Care Acuity Index was a reliable and valid instrument for measurement of acuity in the long-term care hospitalized patient. Inferential statistics also included an analysis of variance used to compare group means and a multiple comparison procedure used to determine the significant differences in group means.

The minimum Cronbach's coefficient alpha reliability for measurement of internal consistency was set at 0.70. This coefficient value was considered to be an acceptable level of reliability for a new instrument.

Content validity was assessed using the diagnostic content validation (DCV) model (Fehring, 1986) in which experts quantify the relevance of each item on the tool and the adequacy with which the tool represents the concept being studied. The process of content validation as outlined by Fehring was as follows:

1. Individuals deemed as experts in the field of long-term acute care reviewed each of the items in the LTA-Care Acuity Index. Each of the items was scored on an ordinal level scale of 1 to 5. Scoring was 1 = not at all characteristic of the long-term acute care patient; 2 = very little characteristics or indicative of the long-term acute care patient; 3 = somewhat characteristic of the long-term acute care patient; 4 = considerably characteristic of the long-

term acute care patient; and 5 = very characteristic of the long-term acute care patient.

2. Weighted ratios were calculated for each of the tool's characteristics. The weights were as follows: 1 = 0; 2 = 0.25; 3 = 0.50, 4 = 0.75; and 5 = 1.00. The range of scores for each of the items for each member of the expert panel was 0-1 so that a value was not given to any item which a member of the expert panel deemed irreverent to the context of the acuity tool.
3. A total DCV score was determined by summing the individual ratio scores within each section of the acuity tool and dividing by the total number of defining characteristics within each of the section.
4. Using the completed evaluation forms, an average score was calculated for each of the instrument's items. Individual items averaging scores of less than 0.80 were classified as minor characteristics, while items with scores equal to or greater than 0.80 were considered major characteristics of acuity.
5. Any item with a combined average content score of less than 0.25 was reviewed for its relationship to the construct.
6. An average coefficient greater than or equal to 0.80 was considered as acceptable estimate of content validity.

Construct validity was assessed by factor analysis using SPSS® Version 9.0 for Windows© to explore the interrelationships among the variables. Factor I represented the pulmonary system; Factor II the cardiovascular system; Factor III the

gastrointestinal system; Factor IV the renal system; Factor V the integumentary system; Factor VI the neurological system; Factor VII medication intervention; Factor VIII intravenous interventions; Factor IX daily activities; Factor X special procedures; and Factor XI the psychosocial system.

Exploratory or principal components factor analysis was used to generate eigenvalues, statistics that represent the relative importance of the factor. A level of 1.00 or higher is a commonly held acceptance criterion for eigenvalues (Ferketich & Muller, 1990). The first factor extracted accounts for the greatest amount of common variance among the variables (Rust & Golombok, 1989). Factors continue to be extracted until a preset eigenvalue of < 0.99 is reached.

Orthogonal varimax factor analysis was performed after the initial principal component analysis. Varimax rotated the factors until the best fit or separation of factors was obtained. The rotation reflected the lowest possible correlation between factors. At least three items loading at 0.30 or two loading with a difference of 0.20 between loadings are necessary for retention of extracted factors (Kline, 1994). Instrument items that result in a factor loading of less than .30 are generally considered to be unclear or ambiguous and may need to be re-written and re-submitted for testing and analysis (Ferketich & Muller, 1990).

A one-way ANOVA was used to test differences in group means for medications, activity point, and total acuity points. Patient care areas were coded as follows: Intensive Care Unit (ICU) = 1, Intermediate Care Unit (IMU) = 2, 3rd

floor general medical-surgical unit = 3, and 4th floor general medical-surgical unit = 4. A Tukey's post hoc test to determine group differences was done for variables with significance levels $\leq .05$.

Summary

The stages of instrument development and analysis were based upon the principles of psychometric theory, beginning with construct definition and concluding with item generation and scaling. The judgment or quantification phase of the study included the reliability and validity testing of the LTA-Care Acuity Index. Internal consistency reliability was reported as Cronbach's alpha. Nurses deemed knowledgeable in the field of long-term acute care served as judges of the instrument's content validity. Factor analysis explored the relationships between the variables. The data were summarized from descriptive and inferential statistical techniques.

It was the intent that the information generated from this study will assist in providing a reliable and valid tool for the measurement of acuity within the long-term acute care hospitalized setting. A contextual, exploratory, descriptive study was undertaken to describe the development validation of the patient acuity instrument, LTA-Care Acuity Index, for long-term acute care patients. The judgment or quantification phase of the study included the reliability and validity testing of the instrument. Reliability was tested using the alpha model for internal consistency; validity testing used the DCV model for content validity. Construct validity was measured through known groups technique and factor analysis.

CHAPTER 4

ANALYSIS OF DATA

The purpose of this study was to examine the reliability and validity of the Long-Term Acute Care Acuity Index (LTA-Care) as an objective measurement of acuity in long-term, acute care, hospitalized patients. Descriptive statistics were used to summarize the characteristics of the sample. Inferential statistics evaluated the tool for its consistency, stability, and relevance to the construct.

An exploratory, nonexperimental, methodological study was undertaken to describe the validation process of the patient classification instrument (LTA-Care Acuity Index) for long-term acute care patients. Before statistical analysis of the LTA-Care Acuity Index, the construct was defined and operationalized through a comprehensive literature review. The psychometric properties of the tool were evaluated in terms of its internal consistency reliability and validity. Construct validity to examine the relationship among variables was determined through known groups technique and factor analysis.

Description of Sample

The sample consisted of 200 subjects consecutively selected from one of four units in a long-term acute care hospital. All study participants were adults, admitted to the hospital for a variety of medical diagnoses requiring long-term, acute inpatient

care. In all cases, the second administration of the tool was used for data analysis. The first administration captures values associated with admission and does not reflect patient acuity scores as they relate to the patient's underlying illness. Acuity measurements were recorded on an identical version of the tool in each of the four patient care areas (ICU, IMU, 3rd, and 4th floors).

A nonprobability convenience sampling technique was used for data collection. Random selection was not feasible for this study due to the large sample base required for instrument testing. Consecutive sampling was appropriate, as there were no criteria regarding patient exclusion. Sample size was determined from power analysis tables for correlation coefficients.

Data analysis was done on individual acuity instruments completed and collected on patients over a 6-month period from 01/01/00 through 06/30/00. To ensure equal representation from each of the care areas, approximately the same number of completed instruments was used for statistical reporting (Table 2).

The hypothesis for this study was formulated from classical measurement theory: The LTA-Care Acuity Index will be a reliable and valid instrument for measuring acuity in an inpatient population: (a) the reliability coefficient alpha will be at least 0.70; (b) the content validation index will be at least 0.80 using the diagnostic content validation model, and (c) construct validity will be established using known groups techniques and factor analysis.

Table 2

Descriptive Statistics: Frequency and Percent of Sample Size

Group	Frequency	Percentage
Intensive Care Unit (ICU)	52	26.0
Intermediate Care Unit (IMU)	48	24.0
3rd Floor Medical-Surgical Unit	49	24.5
4th Floor Medical-Surgical Unit	51	25.5
Total	200	100.0

Findings

Descriptive statistics were used to measure differences in acuity scores between floors. The number of acuity points was determined by totaling all of the activity points for the instrument and adding the sum of the nursing administered medications that the patient received over 24 hours. Acuity scores by location were evaluated for differences based upon total activity scores, total medication scores, and total acuity scores to determine if there were location differences.

Acuity scores were highest among the ICU patients, while the 3rd floor medical-surgical area had the lowest acuity. The difference in total acuity points can be seen in the differences in the average activity points, as there appeared little variation in the average number of medication administrations per floor (Table 3).

Table 3

Descriptive Statistics: Average Medication, Activity, and Acuity Points by Floor

Group	Average Medication Points ± SD	Average Activity Points ± SD	Average Acuity Points ± SD
Intensive Care Unit (ICU)	24.32 ± 8.37	59.46 ± 13.38	83.83 ± 16.09
Intermediate Care Unit (IMU)	25.88 ± 8.64	53.25 ± 13.71	78.23 ± 15.35
3rd Floor Medical-Surgical Unit	24.12 ± 8.53	34.69 ± 15.82	58.94 ± 18.42
4th Floor Medical-Surgical Unit	26.71 ± 9.50	42.90 ± 13.00	69.22 ± 15.39
Total	25.25 ± 8.77	47.68 ± 16.85	72.55 ± 18.72

Reliability

The focus of the first research hypothesis was to determine if the LTA-Care Acuity Index instrument demonstrated internal consistency reliability. The hypothesis was tested using a Cronbach's alpha. The alpha for the entire instrument was 0.7538. Alpha coefficients for each of the instrument's 11 subsets ranged from 0.7237 to 0.7619, thereby suggesting that each of the items equally affected the variation in the total score and deletion of any of the subsets would not significantly improve any of the individual item correlations.

Validity

The diagnostic content validity (DCV) model (Fehring, 1987) was used to examine content validity of the LTA-Care Acuity Index. Ten copies of the acuity tool were sent to registered nurses, experienced in the care of the long-term acute care hospitalized patient. All experts had a minimum of 3 years of long-term health care experience in various capacities ranging from staff level to administrative roles. Each member of the expert panel received a packet with a copy of the acuity tool and written instructions for completion. A self-addressed envelope was included to facilitate return. Nine of the 10 experts responded with completed instrument evaluations. Each defining characteristic was given a weight, ranging from 0-1. Weights were given to each of the items so that the total score could not exceed 1.00. This range of weights further prevented a value to be given to a defining characteristic that was judged by the experts not to be indicative of patient acuity.

Each completed content validation tool was scored by summing the individual ratio scores and dividing by the total number of defining characteristics contained within each of the 77 variables. Variables obtaining an average ratio of 0.80 or greater were considered major characteristics; ratios between 0.50 and 0.80 were considered minor characteristics (Appendix D). Defining characteristics with ratios less than or equal to 0.50 were not included in the total score. The DCV for each variable fell between 0.78 and 0.94, which supported the hypothesis that the instrument is a valid tool for measuring acuity in the long-term in-patient population (Table 4).

Table 4

Results of Content Validation Using the DCV Model

Subset	N	Mean	Min	Max	SD
Pulmonary	9	0.78	0.50	1.00	0.18
Cardiovascular	6	0.81	0.67	0.94	0.09
Gastrointestinal	6	0.81	0.72	1.00	0.11
Renal/Genitourinary	5	0.85	0.69	1.00	0.12
Skin/Wound Care	4	0.91	0.83	0.97	0.06
Neurological	5	0.79	0.69	0.94	0.11
Medications	4	0.90	0.89	0.92	0.02
IVs/Lines	12	0.79	0.61	0.94	0.10
Daily Activities	17	0.92	0.72	1.00	0.08
Special Procedures	6	0.79	0.72	0.86	0.06
Psychosocial	3	0.94	0.86	1.00	0.07

Note. DCV Average = 0.8467 ± 1.272

Construct Validity

A one-way analysis of variance was used to explore difference in group means for the dependent variables: medication administration, activity, and acuity points by the independent variable: location. The level of significance was set at 0.05. It was

predicted that patients in the ICU/IMU setting would have higher acuity scores while patients on the medical-surgical units would have lower acuity scores.

There was no significant difference in total number of medication administrations by unit ($p \geq 0.05$). There was a significant difference in total number of activity and acuity points by unit (Table 5).

Table 5

Analysis of Variance (ANOVA)

	df	Mean Square	F	p
Total Number of Medical Administrations				
Between Groups	3	77.82	1.01	.389
Within Groups	196	76.95		
Total	199			
Total Number of Activity Points				
Between Groups	3	6044.89	30.84	.000
Within Groups	196	196.04		
Total	199			
Total Number of Acuity Points				
Between Groups	3	5765.79	21.56	.000
Within Groups	196	267.45		
Total	199			

A Tukey's multiple comparison test was used to determine where the significant differences in activity and acuity means occurred. Floors were coded as

follows: 1.00 = ICU, 2.00 = IMU, 3.00 = 3rd floor, and 4.00 = 4th floor. Results of the multiple comparison test (Table 6) indicated no significant differences in activity points between the ICU and the IMU. There were significant differences in activity points between ICU and IMU and the 3rd and 4th floors ($p \leq 0.01$). There were no significant differences in activity points between the 3rd and 4th floors.

Table 6

Tukey's Post Hoc Test

Total Activity Points			Total Acuity Points		
(I) Location	(J) Location	p	(I) Location	(J) Location	p
1.00	2.00	.119	1.00	2.00	.393
	3.00	.000		3.00	.000
	4.00	.000		4.00	.000
2.00	1.00	.119	2.00	1.00	.393
	3.00	.000		3.00	.000
	4.00	.001		4.00	.031
3.00	1.00	.000	3.00	1.00	.000
	2.00	.000		2.00	.000
	4.00	.081		4.00	.090
4.00	1.00	.000	4.00	1.00	.000
	2.00	.001		2.00	.031
	3.00	.001		3.00	.090

Differences were also found in total acuity points by floor. There were no significant differences in total acuity points between the ICU and IMU patient care areas. However, there were significant differences in acuity scores between ICU, IMU, and the 3rd and 4th floors, respectively. There were no significant differences in acuity scores between the two general medical-surgical areas.

Factor Analysis Construct Validity

The purpose of doing factor analysis was to explore the underlying relationships between the variables contained in the LTA-Care Acuity Index. The principal component form of factor analysis with a varimax rotation was done using the 200-item database. The maximum likelihood extraction method of principal component factor analysis yielded four factor trends with eigenvalues of 1.00 or greater (Table 7). These four trends accounted for 57% of the variance for the sample. The four factors were then rotated using orthogonal varimax rotation to separate out the factors.

Six of the 11 subscales loaded on Factor I (55%) with a loading value of 0.30 or greater (Table 8). Three of 11 subscales (27%) loaded on Factor II with loading values of 0.30 or greater. Two subscales loaded on Factor III (18%), while three subscales loaded on Factor IV (27%). No subscales had greater than two significant factor loadings.

Table 7

Principal Components Analysis of the LTA-Care Acuity Index Scores

Factor Trends	Eigenvalue * = ≥ 1.00	Proportion Variance	Cumulative Percent
1	4.526	.32	.32
2	1.454	.10	.43
3	1.189	.08	.51
4	1.052	.08	.58
5	1.012	.07	.66
6	0.931	.06	.73
7	0.829	.06	.78
8	0.735	.05	.84
9	0.621	.04	.88
10	0.603	.04	.93
11	0.523	.04	.96
12	0.429	.03	.98
13	<.01	<.01	.99
14	<.01	<.01	1.00

* = Factor Trend

It appears that Factor I represents nursing interventions associated with the patient's clinical condition as the variables contained within these subsets are related to the patient's medical complexity. The integumentary system loaded exclusively on Factor III. Elements contained within the system address compromised skin integrity and take into account nursing time spent on wound care, specifically dressing changes. Special Procedures loaded on Factor IV as did psychosocial system points. Elements

Table 8

Results of Varimax Orthogonal Rotation of Factor Trends

Factor Trends	Factor 1	Factor 2	Factor 3	Factor 4
Pulmonary	.623			
Cardiovascular	.700			
Gastrointestinal	.504			.545
Renal	.677			
Integumentary			.838	
Neurological	.708			
Medications		.700		
Intravenous	.591	.408		
Daily Activities	.551			.540
Special Procedures				.375
Psychological System				.834

contained under special procedures describe interdisciplinary interventions with an emphasis on pre-procedure teaching and post-procedure monitoring. Psychosocial system points take into account other interdisciplinary events including code blue, frequent call light use, and frequent family intervention. The gastrointestinal system

and daily activities loaded on both Factor I and Factor IV. It is possible that there is some ambiguity contained within these subscales as interventions include dependent ADLs, which could be associated with the patient's underlying clinical condition as well as for interdisciplinary care needs. In addition, tube feedings, a variable listed under the gastrointestinal subscale, may be more appropriately placed under daily activities with similar variables that address patients' nutritional needs. The factor trends are summarized in Figure 3.

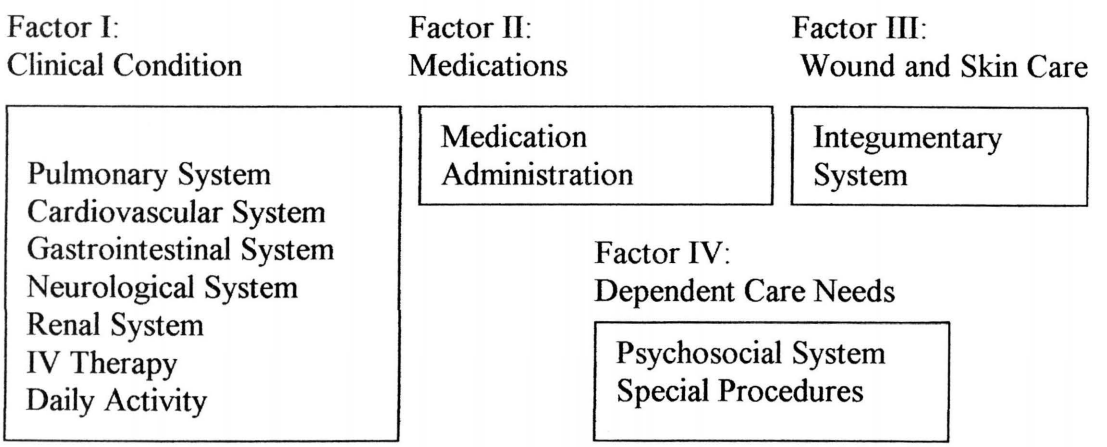

Figure 3. Summary of Factor Trends

Summary of Findings

A total of 200 acuity tools were used to measure the reliability and validity of the LTA-Care Acuity Index for use in the long-term hospitalized adult patient. Data were used in aggregate to measure internal consistency reliability, content validity,

and factor analysis construct validity. Independent measures of medication administration, activity scores, and total acuity scores were analyzed for their construct validity by the dependent variable, location.

A Cronbach's alpha coefficient of 0.75 supported the reliability of the instrument. Each of the elements contained in the instrument had a calculated alpha of greater than 0.70, with a range of 0.04, which indicated that no one element contributed negatively to the overall reliability coefficient.

The calculated DCV index was 0.84, which supported the content validity of the instrument. Each factor had a value of 0.80 or greater. The DCV model allowed for recognition and separation of major and minor characteristics of the acuity tool (Appendix D). In terms of the contribution of each of the categorical variables, each appeared to be a major characteristic of the instrument.

A one-way analysis of variance identified significant between group differences in activity points and acuity points by location. Higher acuity scores were noted in the more acute areas of the hospital (ICU and IMU), while lower acuity scores were found on the general medical surgical units. Variations in activity points contributed to the significant differences in acuity scores by location. Average medication administration did not vary by unit.

A principal component factor analysis was done to examine the interrelationship among the categorical variables. The analysis extracted a four-factor trend. Factor I appears to be related to the patient's overall clinical condition.

Medications loaded exclusively on Factor II. Wound and skin care was represented by Factor III. The gastrointestinal system loaded on Factors I and IV, along with daily activities, which suggested that this factor represents dependent care needs. The variables contained in the gastrointestinal system variable may need to be revised as they appear to not exclusively correlate with clinical condition. Special procedures as well as the psychosocial system loaded on Factor IV, which suggested a relationship of these subscales in terms of interdisciplinary care.

CHAPTER 5

SUMMARY OF THE STUDY

The purpose of this study was to examine the reliability and validity of the LTA-Care Acuity Index, an instrument designed to measure acuity in the long-term adult hospitalized patient. Before the implementation of the acuity tool in this specialized patient population, it was necessary to evaluate the psychometric properties of the instrument.

The underlying assumptions of this study were derived from classical measurement theory. Classical measurement theory also served as a guide in the implementation and evaluation of the research process. The basic principle of classical measurement theory is that the observed (obtained) score is the sum of the true score and the error score. It is also apparent from these assumptions that increased measurement error will result in decreased reliability of the measurement tool. The effect of random error also may be negated and the true score more accurately assessed by determining the average of many independent measures (Nunnally, 1959).

This chapter provides a summary of the purpose for this study and the methods used in conducting the research. In addition, conclusions are offered as well as implications regarding how this research will impact the usefulness of this acuity tool

in the desired in-patient setting. Recommendations for future research also are presented.

Discussion of Findings

Nurses comprise the single largest group of hospital employees while accounting for 25% to 30% of total institutional costs. As the major user of personnel, nursing services must be able to justify their needs, productivity levels, and staff expenses. An ongoing concern of nursing administrators is the appropriate allocation of staffing resources (Alward, 1983). The challenge is not limited to the number of nurses that an administrator should employ, but it also includes a determination of the proper number of staff with the appropriate qualifications to meet patient care requirements an a daily basis.

Hospitals without classification systems often resort to a staffing matrix to determine the allowable number of nursing hours per patient day. This staffing matrix is usually directed by hospital administration and is based on annual budget guidelines which are sometimes determined by non-clinical personnel. Other facilities may use patient-to-nurse ratios as a measure of nurse allocation. Often, these methods are used without consideration of the illness characteristics or psychosocial needs of the patient population.

There are many inconsistencies regarding the number of patient classification systems that are in existence or in use in any type of hospital setting. The most recent

statistics indicate that there may be over 1,000 types of acuity systems being used in hospitals nationwide. In one survey, over 40% of hospitals used a tool that was developed internally (Nagaprasanna, 1988).

A review of the literature supported the account of the variety of classification instruments as they were used for measurement of acuity in different hospital settings and patient populations. In reviewing the literature, a classification tool for measuring acuity in the long-term acute care setting could not be found. In addition, it has been noted in several studies that patient classification systems are not designed to be generic enough that they can be considered universal in nature for use in any type of patient care setting. The type of acuity tool was needed to meet regulatory requirements, to ensure a safe level of testing based upon some measurable parameter, and to provide justification to hospital administration regarding staffing ratios in this patient care setting.

The issues of reliability and validity of patient classification systems are extremely important. As with any instrument or measuring device, some estimate of both reliability and validity must be established before the instrument can be used with confidence. The purpose of this research project was to examine the psychometric properties of the LTA-Care Acuity Index. Reliability, validity, and factor analysis are discussed within the context of classical measurement theory used as the theoretical framework guiding study design and implementation.

The reliability of measures is dependent on the correlation of the sample of variables with the true scores and reflects the degree to which an instrument is free from errors of measurement. Reliability coefficients range from 0.00 to 1.00, with a higher reliability coefficient indicating a stronger relationship. However, if the coefficient values are 1.00, each item on the instrument would be measuring exactly the same thing. A coefficient between 0.70 and 0.80 indicates that an instrument is capable of detecting fine distinctions between levels of the construct (Burns & Grove, 1993).

Validity of an instrument is the extent to which the instrument actually measures the construct being examined, or in this case, the validation of the patient classification instruments in measuring acuity. Content validity indexes reflect the classification system's ability to adequately represent the domain it is supposed to measure, such as patients' requirements for nursing care time (Giovanetti, 1979). Content validity has no empirical basis and relies generally on judgement from content experts.

The results of this study indicated that there was a difference in acuity scores based upon patient location. The ICU, traditionally an area of high acuity in any inpatient population, generated the highest average acuity points. This unit was followed by the intermediate care unit, which in most facilities is a step-down unit that differs in terms of bedside monitoring, equipment, frequency of vital signs, and overall patient stability. The general medical-surgical areas had the lowest average

acuity points. Although they differed in average total acuity points, the principal difference between the two medical-surgical areas was in the average activity points. Values regarding the number of medication administrations compared similarly between floors. There were no significant differences in average acuity scores between the 3rd and 4th general medical-surgical floors. Differences in total acuity averages were seen in the average number of activity points, which took into consideration the severity of illness as well as the frequency and amount of nursing time required to meet patient care needs.

There appears to be a relationship between the acuity scores and staffing levels in each of the patient care areas. Traditionally, the more acutely ill the patient, the higher the staffing levels. Intensive care units may staff one nurse for every two patients, while each nurse on the general medical-surgical units may have five or more patients.

Reliability

Because the LTA-Care Acuity Index is a newly developed instrument, there are no previously reported estimates of reliability. Internal consistency reliability of the instrument using Cronbach's alpha was established, giving some assurance that the reliability coefficient obtained on the LTA-Care Acuity Index may be replicated if used again in the long-term acute care patient population. The internal consistency reliability alpha of 0.7538 met the recommended minimal criteria of 0.70 (Brink &

Wood, 1993) for a newly developed instrument. Failing to assess internal consistency makes it difficult for the user to determine the sensitivity of the measure and where all of the variables contribute to the measurement of acuity. Findings from this study demonstrated that the instrument had internal consistency reliability.

Content Validity

Validity is considered a more important evaluative measure than reliability, although it is addressed in the literature with less frequency (Knapp, 1985). For an instrument to be valid, it must be reliable; however, it can be reliable without being valid. If the instrument demonstrates validity, then it is measuring the construct it is supposed to be measuring. Unlike reliability, validity is a stable property of an instrument as long as the instrument is used in the same manner for which it was developed.

When using an instrument in a new or adapted way, content validity, or the extent to which the instrument covers a constant domain, should be evaluated (Knapp, 1985). Because the LTA-Care Acuity Index was adapted with some minor revisions from the Therapeutic Intervention Scoring System, content validity was measured. The method chosen for determining content validity was the diagnostic content validity (DCV) model (Fehring, 1987). Nine content experts evaluated the LTA-Care Acuity Index for the degree in which the subsets in the instrument adequately represented the construct of acuity. The calculated DCV was 0.84, giving support to the assumption

that the variables contained in the instrument measured acuity in the long-term hospitalized patient. Each of the domains within the LTA-Care Acuity Index fell within the established measure of acceptance as contributors to the construct. Major and minor characteristics were extracted from the analysis. All of the 76 variables contained in the instrument were deemed to be characteristics that influenced acuity in the long-term, acute care hospitalized patient.

Construct Validity

One of the most efficient ways to establish construct validity is through use to the known groups approach to assess the extent to which an instrument measures the intended concept (Burns & Grove, 1993). Construct validity can only be established over time through multiple methods and a diversity of subjects. Construct validity identifies valid physiological constructs that allow for evaluation of known group characteristics as well as identification of the interrelationship among the individual subsets.

Exploratory factor analysis sorts variables under conceptual factors and is used to revise the instrument with recognition of the strength of the loadings on the extracted factors. Factor analysis studies how individual items interrelate. The sum of squares of the factor loadings of each factor reflects the proportion of variance explained by each factor (Kline, 1994). Four factor trends (or eigenvalues) accounted for 57% of the variance. Each of the subscales in the LTA-Care Acuity Index loaded

on one or more of the four factors with a coefficient value of 0.30 or greater. Four of the five physiological systems-based variables loaded exclusively on the first factor. Medications loaded exclusively on Factor II. Nursing interventions associated with this subscale are differentiated from the category of total medication administrations as these subscale variables are primarily related to PRN as opposed to scheduled medications. The integumentary system loaded exclusively on Factor III. This outcome may be correlated to the elements contained within this variable as they relate to nursing interventions regarding wound care and maintenance of skin integrity. A review of the variables contained in the gastrointestinal system will be evaluated, as this system loaded equally on Factors I and IV. This outcome possibly may be due to the influence of tube feeding scores contained in the gastrointestinal subset. This variable may be more appropriate in the daily activities subscale.

When developed and used appropriately, patient classification systems allow for efficient and effective management of nursing resources. Benefits have been derived from such use, not only in short-term, daily allocation of staff, but also in budget planning. The management of these resources will aid in decision-making regarding care priorities. Using a scientific basis for staffing plans and allotment of personnel on a daily basis can result in improved nursing care and appropriate utilization of resources. All of these factors can contribute to both job satisfaction for the providers of patient care as well as patient satisfaction with the delivery of care.

Conclusions

Based on the results of the psychometric testing of the LTA-Care Acuity Index, the following conclusions are offered:

1. The LTA-Care Acuity Index demonstrated acceptable levels of internal consistency reliability for the measurement of acuity in the long-term acute care hospital setting.
2. The LTA-Care Acuity Index demonstrated acceptable levels of content validity for this sample of hospitalized patients.
3. There were significant known group differences in acuity scores by patient care areas. The ICU had the highest acuity scores, followed by the IMU and the general medical-surgical areas.
4. Differences in acuity scores were attributed to variations in activity points. Average medication administrations did not significantly differ by location.

Implications

The implications of this study on professional nursing and on the care of the long-term hospitalized patient are as follows:

1. The findings of this study have demonstrated that it is possible to develop and utilize a patient acuity index that has reliability and validity in a long-term care setting. With the staffing, utilization, and reimbursement problems in

long-term care, these findings are of great importance to nurse administrators and nurses working in this patient care setting.

2. Acuity tools may be used as an adjunct to locating areas within the hospitals in which nursing resources are inadequate and need additional support. These tools also can help identify areas where staffing is sufficient or overstaffing is present.
3. The utilization of a reliable and valid tool for measuring acuity has the potential to assist in the prioritization of nursing care.
4. The complexity of patient care activities can be quantitatively evaluated with the use of a patient acuity tool.
5. A valid and reliable acuity indicator may be used as partial criteria for the assignment of patients to different levels of care.
6. Having a reliable and valid system to measure individual acuities may impact the type of care that is needed by individuals, not by diagnoses.

Recommendations for Further Research

The following recommendations for further study are offered based upon results of this research:

1. Development of a method to factor in the effects of the physical layout of the facility or the distance between patient rooms to determine the amount of

nonproductive time spent going from one location to the next would enhance the effectiveness of using a long-term care acuity instrument.

2. Using modern computer technology which can be updated constantly, a means needs to be developed to determine any differences in patient acuity by calculating acuity scores by shift or by hour. This capability would enable nursing administration to justify any changes in staffing levels for different shifts.

3. This study should be replicated in another long-term acute care environment with more diverse clientele, using different forms of reliability and validity testing. This effort could provide further evidence of flexibility, reliability, and validity of the instrument.

4. A taxonomy of categories and/or classes of patients based upon acuity scores should be developed for long-term care.

5. A study should be designed using patient acuity scores as a guide for patient assignments as opposed to arbitrarily assigning a nurse to a patient load because that number of patients is the accepted hospital standard. Outcomes measurement could assess the effectiveness of this type of staffing plan.

6. Acuity values and levels should be factored into the annual nursing budget.

7. Further investigation should occur into the use of acuity scores as partial criteria for appropriateness of patient discharge to home or to another facility that provides a different level of care.

8. Investigation should occur to study the development of a mechanism to incorporate acuity scores into utilization review, which would evolve into a method to compare acuities with the cost of hospitalization.
9. A study should be designed to correlate various outcome measures with acuity scores to determine the predictive value of the instrument.
10. A theoretical model needs to be developed for the psychometric testing of patient classification systems.
11. Studies to find ways to incorporate the patient classification system into the hospital's electronic medical record are vital. A quality assurance system that would facilitate ongoing monitoring and trending of acuity scores should complement this activity.

While patient classification systems are not without flaws in determining the true needs of patients, when used appropriately, they do provide a more rational approach to the problem of nurse staffing. This study has demonstrated that reliability and validity of acuity instruments can be tested so that nurses, hospital administrators, funders, and the public can have increasing confidence in the acuity systems in place in long-term care facilities.

REFERENCES

Alward, R. R. (1983). Patient classification systems: The ideal vs. reality. Journal of Nursing Administration, 13(2), 14-19.

Andro, R. J., Robertson S., & Glandon, G. L. (1987). Financial implications of an acuity-based periodic staff adjustment model. Nursing Management, 18(10), 21-25.

Aronow, D. B. (1988). Severity of illness measurement: Applications in quality assurance and utilization review. Medical Care Review 45, 339-366.

Aydelotte, M. K. (1973). Standard I: Staffing for quality care. Journal of Nursing Administration 3(2), 33-36.

Batty, K., Mooney, M., & Lowry, C. (1990). Patient classification: A visible difference. Nursing Management, 21(9), 71-75.

Bermas, N. F., & Van Slyck, A. (1984, November 14). Patient classification systems and the nursing department. Hospitals, 99-100.

Brink, P. J., & Wood, M. J. (1989). Advanced design in nursing research. Newbury Park, CA: Sage.

Buckle, J. M., Horn, S. D., & Simpson, R. L. (1991). Nursing care classification: A conceptual model. Applied Nursing Research, 4, 100-106.

Burns, N. (1987). The practice of nursing research: Conduct, critique and utilization. Philadelphia, PA: Saunders.

Burns, N., & Grove, S. I. (1993). The practice of nursing research (2nd ed.). Philadelphia, PA: Saunders.

Carmines, E. G., & Zeller, R. A. (1979). Reliability and validity assessment. Newbury Park: Sage.

Carr-Hill, R. A., & Jenkins-Clarke, S. (1995). Measurement systems in principle and in practice: The example of nursing workload. Journal of Advanced Nursing, 22, 221-225.

Clermont, G., Angus, D. C., Linde-Zwirble, W. T., Lave, J. R., & Pinsky, M. R. (1998). Measuring resource use in the ICU with computerized therapeutic intervention scoring system-based data. Chest, 113, 434-442.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Cullen, D. J., Civetta, J. M., Briggs, B. A., & Ferara, L. C. (1974). Therapeutic intervention scoring system: A method for quantitative comparison of patient care. Critical Care Medicine, 2, 57-60.

Cullen, D. J., Nemeskal, A. R., & Zaslavsky, A. M. (1994). Intermediate TISS: A new therapeutic intervention scoring system for non-ICU patients. Critical Care Medicine, 22, 1406-1411.

De Groot, H. A. (1989a). Patient classification system evaluation: Part one: Essential system elements. Journal of Nursing Administration, 19(6), 30-35.

De Groot, H. A. (1989b). Patient classification system evaluation: Part two: System selection and implementation. Journal of Nursing Administration, 19(7), 24-27.

Des Ormeaux, S. P. (1977). Implementation of the C.A.S.H. patient classification system for staffing determination. Supervisor Nurse, 8(4), 29-35.

Dijkers, M., & Paradise, T. (1986). PCS: One system for both staffing and costing. Nursing Management, 17(1), 25-34.

Ebner, M. K. (1985). Reliability and validity basics for evaluation classification systems. Nursing Economics, 3, 324-328.

Fehring, R. (1987). Method to validate nursing diagnosis. Heart and Lung, 16, 625-629.

Ferketich, S. (1991). Focus on psychometrics: Aspects of item analysis. Research in Nursing and Health, 14, 165-168.

Ferketich, S., & Muller, M. E. (1990). Factor analysis revisited. Nursing Research, 39, 59-62.

Georgette, J. K. (1970). Staffing by patient classification. Nursing Clinics of North America, 5, 329-339.

Giovanetti, P. (1979). Understanding patient classification systems. Journal of Nursing Administration, 9(2), 4-9.

Giovanetti, P., & Mayer, G. G. (1984). Building confidence in patient classification systems. Nursing Management, 15(8), 31-34.

Haas, S. A. (1988). Patient classification systems: A self-fulfilling prophecy. Nursing Management, 19(5), 56-62.

Iezzoni, L. I. (1990). Severity of illness measures: Comments and critiques. Medical Care, 28, 757-761.

Jackson, B. S., & Resnick, J. R. (1982). Comparing classification systems. Nursing Management, 13(2), 13-19.

Joint Commission on Accreditation of Healthcare Organizations. (1998). Accreditation manual for hospitals, Vol. 1: Standards. Oakbrook Terrace, IL: The author.

Keene, A. R., & Cullen, D. (1983). Therapeutic intervention scoring system: Update 1983. Critical Care Medicine, 11, 1-3.

Kim, J. O., & Mueller, C. W. (1978). Introduction to factor analysis: What it is and how to do it. Newbury Park, CA: Sage.

Kline, P. (1994). An easy guide to factor analysis. New York: Routledge.

Knapp, T. R. (1985). Validity, reliability and neither. Nursing Research, 34, 189-192.

Lewis, N. L., & Carini, P. V. (1984). Nursing staffing and patient classification. Rockville, MD: Aspen.

LoBiondo-Wood, G., & Haber, J. (1994). Nursing research: Methods, critical appraisal & utilization. St. Louis: Mosby.

Lynn, M. R. (1986). Determination and quantification of content validity. Nursing Research, 35, 382-385.

Lynn, M. R. (1989). Instrument reliability and validity: How much needs to be published? Heart and Lung, 18, 421-423.

Malloch, K., & Conovaloff, A. (1999). Patient classification systems (1). Journal of Nursing Administration, 29(7-8), 49-56.

Medicus System Corporation. (1983). Medicus nursing productivity and quality system: Overview. Evanston, IL: The Author.

Moreno, R., & Morais, P. (1997). Validation of the simplified therapeutic intervention scoring system on an independent database. Intensive Care Medicine, 23, 640-644.

Nagaprasanna, B. R. (1988). Patient classification systems: Strategies for the 1990's. Nursing Management, 19, 105-112.

Niemeier, D. F., & Reed, G. B. (1985). A monitoring system for patient classification in critical care. Dimensions of Critical Care Nursing, 4, 110-118.

Norusis, M. J. (1990). SPSS base system's user guide. Chicago: SPSS.

Nunnally, J. C. (1959). Tests and measurements: Assessment and prediction. New York: McGraw-Hill.

Nunnally, J. C. (1967). Psychometric theory. New York: McGraw-Hill.

Nunnally, J. C. (1970). Introduction to psychological measurement. New York: McGraw-Hill.

Nunnally, J. C. (1978). Psychometric theory. New York: McGraw-Hill.

Pedazur, E. J., & Schmelkin, L. P. (1991). Measurement, design, and analysis: An integrated approach. Hillsdale, NJ: Lawrence Erlbaum.

Phillips, C. Y., Castorr, A., Prescott, P. A., & Soeken, K. (1992). Nursing intensity: Going beyond patient classification. Journal of Nursing Administration, 22 (4), 46-52.

Polit, D. F., & Hungler, B. P. (1987). Nursing research: Principles and methods (3rd ed.). Philadelphia, PA: Lippincott.

Portney, L. G., & Watkins, M. P. (2000). Foundations of clinical research: Applications to practice (2nd ed.). Upper Saddle River, NJ: Prentice Hall Health.

Reinert, P., & Grant, D. R. (1981). A classification system to meet today's needs. Journal of Nursing Administration, 11(1), 21-25.

Reis Miranda, D., De Rijk, A., & Schaufeli, W. B. (1996). Simplified therapeutic intervention scoring system: The TISS-28 items--results from a multicenter study. Intensive Care Medicine, 24, 64-73.

Reitz, J. A. (1985). Toward a comprehensive nursing intensity index: Part II, testing. Nursing Management, 16(9), 31-40.

Rust, J., & Golombok, S. (1989). Modern psychometrics. New York: Routledge.

Schroeder, R. E., Rhodes, A. M., & Sheilds, R. E. (1984). Nurse acuity systems: CASH vs. GRASP (a determination of nurse staff requirements). Nursing Forum, 21(2), 72-77.

Scribante, J., Muller, M. E., & Lipman, J. (1996). Development and validation of a critical care patient classification system. American Journal of Critical Care, 5, 282-288.

Shelley, S. (1984). Research methods in nursing and health. Boston: Little, Brown.

Slocumb, E., & Cole, F. (1991). Method for obtaining content validity through the expert panel. Applied Nursing Research, 4, 192-195.

Summers, S. (1993). Establishing the reliability and validity of a new instrument: Pilot testing. Journal of Post-Anesthesia Nursing, 8, 124-127.

Vaughn, R. G., & MacLeod, V. (1980). Nurse staffing studies: No need to reinvent the wheel. Journal of Nursing Administration, 10(3), 9-15.

Wagner, D., & Draper, E. (1984). Acute physiology and chronic health evaluation (APACHE II) and medicare reimbursement. Health Care Financing Review (Suppl.), 91-105.

Wagner, D., Draper, E., & Knaus, W. (1989). APACHE III study design: Analytic plan for evaluation of severity and outcome in intensive care unit patients. Analysis: Quality of care. Critical Care Medicine, 17(12 Pt 2), S-210-S-212.

Wagner, D., Knaus, W., & Draper, E. (1983). Statistical validation of a severity of illness measure. American Journal of Public Health, 73, 878-884.

Waltz, C. F., Strickland, O. L., & Lenz, E. R. (1991). Measurement in nursing research (2nd ed.). Philadelphia: F. A. Davis.

Waltz, C. W., & Bausell, R. B. (1981). Nursing research: Design, statistics and computer analysis. Philadelphia: F.A. Davis.

Whitney, J. D., & Killien, M. G. (1987). Establishing predictive validity of a patient classification system. Nursing Management, 18(5), 84-86.

Williams, M. A. (1988). When you don't develop your own: Validation methods for patient classification systems. Nursing Management 19(3), 90-96.

Zimmerman, J. E., Wagner, D. P., Draper, E. A., Wright, L., Alsola, C., & Knaus, W. A. (1998). Evaluation of acute physiology and chronic health evaluation. III: Predication of hospital mortality in an independent database. Critical Care Medicine, 8, 1317-1326.

APPENDIX A

LONG-TERM CARE ACUITY INDEX (LTA-CARE)

LTA-Care

ACUITY INDEX

Subject ID number:

FLOOR NUMBER											
DATE											
PULMONARY											
Extubation first 24 hours	4	4	4	4	4	4	4	4	4	4	4
Decannulation first 24 hours	3	3	3	3	3	3	3	3	3	3	3
Assisted vent with artificial airway	4	4	4	4	4	4	4	4	4	4	4
Active weaning	4	4	4	4	4	4	4	4	4	4	4
Frequent suction: trach, ETT	3	3	3	3	3	3	3	3	3	3	3
Trach tube	2	2	2	2	2	2	2	2	2	2	2
Chest tube	3	3	3	3	3	3	3	3	3	3	3
Nasal cannula	1	1	1	1	1	1	1	1	1	1	1
Respiratory isolation	5	5	5	5	5	5	5	5	5	5	5
CARDIOVASCULAR											
Telemetry	1	1	1	1	1	1	1	1	1	1	1
VS > Q shift	2	2	2	2	2	2	2	2	2	2	2
Fluid challenges	3	3	3	3	3	3	3	3	3	3	3
Drip for intravascular stability	3	3	3	3	3	3	3	3	3	3	3
Bedside procedure monitoring	3	3	3	3	3	3	3	3	3	3	3
Continuous IV med administration requiring lab &/or VS monitoring	3	3	3	3	3	3	3	3	3	3	3
Initial Digitalization	3	3	3	3	3	3	3	3	3	3	3
GASTROINTESTINAL											
Tube Feedings	2	2	2	2	2	2	2	2	2	2	2
NGT suction	1	1	1	1	1	1	1	1	1	1	1
Assisted bowel evacuation	2	2	2	2	2	2	2	2	2	2	2
Stools > 1/shift	2	2	2	2	2	2	2	2	2	2	2
Diarrhea	3	3	3	3	3	3	3	3	3	3	3
Colostomy/ileostomy	3	3	3	3	3	3	3	3	3	3	3
RENAL/GENTOURINARY											
Incontinent	4	4	4	4	4	4	4	4	4	4	4
Standard I & O	1	1	1	1	1	1	1	1	1	1	1
Foley	1	1	1	1	1	1	1	1	1	1	1
Bladder irrigation	3	3	3	3	3	3	3	3	3	3	3
Hemodialysis	1	1	1	1	1	1	1	1	1	1	1
SKIN/WOUND CARE											
Complex dressing change	4	4	4	4	4	4	4	4	4	4	4
Decubitus dressing (simple)	2	2	2	2	2	2	2	2	2	2	2
Catastrophic wounds	4	4	4	4	4	4	4	4	4	4	4
Replace soiled dressing	1	1	1	1	1	1	1	1	1	1	1
NEUROLOGICAL											
Slightly confused	2	2	2	2	2	2	2	2	2	2	2
Confused/disorientated	4	4	4	4	4	4	4	4	4	4	4
Comatose	3	3	3	3	3	3	3	3	3	3	3
Seizure Precautions	2	2	2	2	2	2	2	2	2	2	2
Restraints (mitts/vest/wrist)	4	4	4	4	4	4	4	4	4	4	4
MEDICATIONS											
IVP meds/ VS documentaion	1	1	1	1	1	1	1	1	1	1	1
PRN meds < 4/shift	2	2	2	2	2	2	2	2	2	2	2
PRN meds > 4/shift	3	3	3	3	3	3	3	3	3	3	3
Total # meds adm./24 hours											

FLOOR NUMBER											
DATE											
IV'S/LINES											
Insulin drip with q 1 hour blood sugars	4	4	4	4	4	4	4	4	4	4	4
Central line/TPN	1	1	1	1	1	1	1	1	1	1	1
Blood sugars q 6 hours with sliding scale	2	2	2	2	2	2	2	2	2	2	2
Blood sugars q 4 hours with sliding scale	3	3	3	3	3	3	3	3	3	3	3
> 2 IV antibiotics	2	2	2	2	2	2	2	2	2	2	2
Routine tubing change	1	1	1	1	1	1	1	1	1	1	1
Continuous sedation drip	2	2	2	2	2	2	2	2	2	2	2
Blood products this shift	3	3	3	3	3	3	3	3	3	3	3
Blood products previous shift	1	1	1	1	1	1	1	1	1	1	1
Bolus IVPB not scheduled	2	2	2	2	2	2	2	2	2	2	2
Peripheral IV saline lock	1	1	1	1	1	1	1	1	1	1	1
Start peripheral IV	2	2	2	2	2	2	2	2	2	2	2
DAILY ACTIVITIES											
Discharge teaching	1	1	1	1	1	1	1	1	1	1	1
Total lift to chair	2	2	2	2	2	2	2	2	2	2	2
Total bed bath	3	3	3	3	3	3	3	3	3	3	3
Completely dependent ADLs	3	3	3	3	3	3	3	3	3	3	3
Daily weight	2	2	2	2	2	2	2	2	2	2	2
Psychological support > 1hour/24 hours	2	2	2	2	2	2	2	2	2	2	2
Pre-op prep and teaching	1	1	1	1	1	1	1	1	1	1	1
Knowledge deficit/med. teaching	1	1	1	1	1	1	1	1	1	1	1
Admission assessment and history	2	2	2	2	2	2	2	2	2	2	2
Care plan initiation	2	2	2	2	2	2	2	2	2	2	2
Care plan modification	2	2	2	2	2	2	2	2	2	2	2
Feeders	4	4	4	4	4	4	4	4	4	4	4
Linen change (1 x # 1of changes)											
Frequent turns and reposition	2	2	2	2	2	2	2	2	2	2	2
Obesity with immobility	3	3	3	3	3	3	3	3	3	3	3
Cont. Nursing time > 60 minutes (4 x # of nurses)											
Cont. Nursing time 30 to 60 minutes (3 x # of nurses)											
SPECIAL PROCEDURES											
Bedside procedures-assist	4	4	4	4	4	4	4	4	4	4	4
Procedures outside of hospital	2	2	2	2	2	2	2	2	2	2	2
Code Blue	4	4	4	4	4	4	4	4	4	4	4
Unplanned transfer to ICU/IMU	4	4	4	4	4	4	4	4	4	4	4
Post-op first 24 hours	2	2	2	2	2	2	2	2	2	2	2
MRSA/VRE isolation	1	1	1	1	1	1	1	1	1	1	1
PSYCHOSOCIAL											
Patient call light > 1/hour	3	3	3	3	3	3	3	3	3	3	3
Imminent patient demise	3	3	3	3	3	3	3	3	3	3	3
Frequent family intervention	4	4	4	4	4	4	4	4	4	4	4
TOTAL + MEDICATIONS											

APPENDIX B

AGENCY APPROVALS

TEXAS WOMAN'S UNIVERSITY
DENTON DALLAS HOUSTON
HUMAN SUBJECTS REVIEW COMMITTEE - HOUSTON CENTER

EXEMPT REVIEW
Application to the Human Subjects Review Committee

This form must be completed if the research committee (for student research) or the department coordinator (for faculty research) decides that the proposed research is exempt from Full Review or Expedited Review by the HSRC. A proposal may be eligible for Exempt Review if any of the following conditions is met:

1) only minimal risk to subjects, as described in the **Human Subjects in Research: Institutional Review Board Policies and Procedures**, pp. 11-12;

   and/or

2) the project will be completed at another institution or in collaboration with investigators at another institution, and that institution's IRB has provided written approval for the proposal as described. **To be eligible for this exemption a signed copy of the institution's current IRB approval form must be attached to this application. If applicable, attach a memo indicating the student's role in the approved study;**

   and/or

3) **the project involves an analysis of a data set generated from a currently approved project.**

For **Exempt Review** by the TWU Human Subjects Review Committee, submit three copies of this form, any relevant Informed Consent Forms, surveys, questionnaires, and (if applicable) the collaborating institution's signed IRB approval form. Approval is required prior to the initiation of the research project. The investigator will be notified if the Human Subjects Review Committee requires additional information.

To complete this form electronically, type information into the blanks provided. If your typing fills the blank, text will wrap automatically. Print out, secure appropriate signatures, and submit three copies (along with accompanying documents) to the Office of Research, MJG 913. Paper-clip each of the copies—no staples, please.

Principal Investigator(s) Kathryn S. Spiegel SS# 460-78-8676

_____ SS# _____

Faculty Advisor (if applicable) K. Lynn Wieck Ph.D. Dept. Nursing

Title of Study

Reliability and Validity Testing of an Acuity Tool for use in the Long-Term Acute Care Hospital Setting

Justification for Exempt Review status Condition 1: The study involves minimal risk to subjects. Subjects will not be required to do anything in terms of participation. All patient identifiers will be blocked out.

Estimated beginning date of the study 02/01/01

Estimated duration of the study 6 months

Research being conducted for (place an X in the appropriate blank):

____ Professional Paper	_X_ Dissertation	____ Pilot Study
____ Thesis	____ Class Project	____ Faculty

Exempt Review Form

HSRC-H 1999-9

Is this research being conducted for a non-university sponsor?

_____ Yes; Name of Sponsor _____

__X__ No

If you are using an electronic form, fill in the blanks provided below. Text will wrap automatically. If you are competing a hardcopy form, attach additional typed page(s) as needed.

1. Give a brief description of the study. Describe the subjects (How many subjects? Are they adults? Minors?) and the procedures that relate to their participation (What will the subjects do? What will be done to them? Where will the study will be conducted?). **(If applicable, identify collaborating institutuion and your role in the study).**

   Data will be collected from a long-term acute care hospital in a large metropolitan area. All patients admitted to the facility are 18 years of age or older.

   The sampling frame will consist of all patients admitted to the long-term acute care hospital between 1/1/00 and 6/1/00. The data extraction and analysis will be done from 2/1/01 to 6/30/01 on previously completed data. The estimated sample size is 150. There are no restrictions regarding patient inclusion, as every patient received initial and weekly acuity scores as a function of routine nursing staff responsibilities.

   This study will use an observer-rated tool. Because the research involves data routinely collected by nursing staff members, written informed consent will not be obtained from the patients. Acuity tools are not part of the patient's permanent medical record and LTA-Care acuity scores will not be used as a provision of patient care. In order to ensure patient anonymity, only copies of completed acuity tools will be provided to the researcher and entered into the database.

   The purpose of the research is to establish the reliability and validity of the LTA-Care acuity tool for adult patients hospitalized in a long-term acute care setting. All patient identifiers (name, room number and date of admission) with the exception of location (2nd, 3rd or 4th floor) will be blacked out on the original acuity tool before photocopying. This process will eliminate any mechanism for associating an acuity score with individual patient identifiers. Only copies of completed acuity tools without patient identifying information will be provided to the researcher for statistical analysis.

2. What are the potential risks to the human subjects involved in this research or investigation?

   Patient confidentiality with the potential for disclosure of the names and their identifying clinical information. To reduce that risk, the researcher will be using regularly collected data without any individual patient identifiers. Additionally, all information will be kept locked when not in use and will be destroyed in 5 years after completion of the study.

SIGNATURE REQUIREMENTS

1. For students

The research protocol and the HSRC application have been read and approved by the members of the student's research committee:

Names of Committee Members	Signatures	Date
K. Lynn Wieck Ph.D	K. Lynn Wieck	1/22/01
Ann T. Malecha Ph.D	Ann T. Malecha	1/23/01

Exempt Review Form

Miguel F. DaCunha Ph.D

2. For faculty

The research protocol and the HSRC application have been read and approved by the academic administrator:

Name of Academic Administrator	Signature	Date

Approved by HSRC Chair Jeberson Zabowski Date 1-25-01

HSRC-H 1999-10

Vencor Hospital • Houston

6441 Main Street
Houston, Texas 77030
(713) 790 • 0500
(713) 790 • 1755 Fax

November 20, 2000

Kathryn S. Spiegel
10727 Shawnbrook
Houston, Texas 77071

Re: Dissertation

Title: RELIABILITY AND VALIDITY TESTING OF AN ACUITY TOOL FOR USE IN THE LONG-TERM ACUTE CARE HOSPITAL SETTING

Vencor Hospital Houston, a Long Term Acute Care Facility, has granted Kathryn S. Spiegel permission to use the patient acuity tool data collected between January 1, 2000 and June 30, 2000, to calculate an estimate of the reliability and validity of the patient acuity instrument. This data is to be used only in aggregate form without either patient or facility identifiers for research purposes and to advance the body of nursing knowledge. Vencor requests access to the data analysis as it regards the reliability and validity of the patient acuity tool.

Sincerely,

Barbara A. Bush

Barbara A. Bush, RN,MSN
Assistant Administrator Clinical Operations

APPENDIX C

CONTENT VALIDATION FORMS FOR EXPERT PANEL

INSTRUCTIONS FOR COMPLETION
CONTENT VALIDATION
LTA-Care ACUITY INDEX

1. The previous page contains a sample of the format design of the acuity tool. The design allows for multiple administrations of the tool for each patient.

2. Each variable has a corresponding number. The number represents the complexity of the task; numbers are correlated to degree of complexity (higher numbers = greater complexity).

3. Calculations performed by nursing personnel are required for elements without corresponding numbers.

4. Medication administration scores represent the number of medications plus the number of administrations of these medications. Medication administration only includes those medications given by nursing personnel.

5. Total acuity score represents the sum of acuity points plus the number of administrations of medications within time frame of 24 hours.

6. Competed validation tools are to be returned in the self-addressed stamped envelope. Please do not include any individual identifiable information, i.e., your name, position, or facility number.

Reliability and Validity Testing of A Patient Acuity System
In the Long-Term Acute Care Hospital Setting

Study Objective: Validate the appropriateness of an instrument to measure acuity on the long-term acute care hospitalized patient.

A weight has been assigned to the following variables based upon the complexity of the item in terms of providing patient care. Please review the following variables and their weights and determine to what extent, in your opinion, each variable is congruent with patient care. Place an "x" in the column corresponding to your answer. Please do not mark more than one column for each of the items.

ITEM	POINT VALUE	Very characteristic of the long-term acute care patient	Considerably characteristic of the long-term acute care patient	Somewhat characteristic of the long-term acute care patient	Very little characteristic of the long-term acute care patient	Not very characteristic of the long-term acute care patient
PULMONARY						
Extubation first 24 hours	4					
Decannulation first 24 hours	3					
Assisted vent with artificial airway	4					
Active weaning	4					
Frequent suction: Trach, ETT	3					
Trach tube	2					
Chest tube	3					
Nasal cannula	1					
Respiratory isolation	5					
CARDIOVASCULAR						
Telemetry	1					
VS > Q shift	2					
Fluid challenges	3					
Continuous IV med administration requiring lab &/or VS monitoring	3					
IVP anti-arrhythmics	3					
Initial digitalization	3					

ITEM	POINT VALUE	Very characteristic of the long-term acute care patient	Considerably characteristic of the long-term acute care patient	Somewhat characteristic of the long-term acute care patient	Very little characteristic of the long-term acute care patient	Not very characteristic of the long-term acute care patient
GASTROINTESTINAL						
Tube feedings	2					
NGT suction	1					
Assisted bowel evacuation	2					
Stools > 1 shift	2					
Diarrhea	3					
Colostomy/Ileostomy	3					
RENAL/GENITOURINARY						
Incontinent	4					
Standard I & O	1					
Foley	1					
Bladder irrigation	3					
Hemodialysis/Peritoneal dialysis	1					
SKIN/WOUND CARE						
Complex dressing change	4					
Decubitus dressing (simple)	2					
Catastrophic wounds	4					
Replace soiled dressing	1					
NEUROLOGICAL						
Slightly confused	2					
Confused/disorientated	4					
Comatose	3					
Seizure precautions	2					
Restraints (mitts/vest/wrist)	4					
MEDICATIONS						
IVP meds/VS documentation	1					
PRN meds < 4 /shift	2					
PRN meds > 4/shift	3					
Total # meds adm. 24 hours						

ITEM	POINT VALUE	Very characteristic of the long-term acute care patient	Considerably characteristic of the long-term acute care patient	Somewhat characteristic of the long-term acute care patient	Very little characteristic of the long-term acute care patient	Not very characteristic of the long-term acute care patient
IV'S/LINES						
Insulin drip with Q 1 hour BS	4					
Central line with TPN	1					
BS q 6 hours with sliding scale	2					
BS q 4 hours with sliding scale	3					
> 2 IV antibiotics	2					
Routine tubing change	1					
Continuous sedation drip	2					
Blood products this shift	3					
Blood products previous shift	1					
Bolus IVPB not scheduled	2					
Peripheral IV saline lock	1					
Start peripheral IV	2					
DAILY ACTIVITIES						
Discharge teaching	1					
Total lift to chair	2					
Total bed bath	3					
Completely dependent ADLs	3					
Daily weight	2					
Psychological support > 1hour/24	2					
Pre-op prep and teaching	1					
Knowledge deficit/med. teaching	1					
Admission assessment & history	2					
Care plan initiation	2					
Care plan modification	2					
Feeder	4					
Linen change (1 x # of changes)						
Frequent turns and reposition	2					
Obesity with immobility	3					

ITEM	POINT VALUE	Very characteristic of the long-term acute care patient	Considerably characteristic of the long-term acute care patient	Somewhat characteristic of the long-term acute care patient	Very little characteristic of the long-term acute care patient	Not very characteristic of the long-term acute care patient
DAILY ACTIVITIES (CONT)						
Continuous nursing time > 60 minutes (4 x # of nurses)						
Continuous nursing time 30-60 minutes (3 x # of nurses)						
SPECIAL PROCEDURES						
Bedside procedures-assist	4					
Procedures outside hospital	2					
Code Blue	4					
Unplanned transfer to IMU/ICU	4					
Post-op first 24 hours	2					
MRSA/VRE isolation	1					
PSYCHOSOCIAL						
Patient call light > 1/hour	3					
Imminent patient demise	3					
Frequent family intervention	4					
TOTAL INTERVENTION POINTS						
TOTAL + MEDICATIONS						

APPENDIX D

TABLE OF VALIDATION DATA AND

DCV SUMMARY

	N	Mean	Min	Max	Std. Dev.	Var.
PULMONARY	9	0.78	0.50	1.00	0.18	0.03
Extubation first 24 hours	9	0.67	0.25	1.00	0.28	0.08
Decannulation first 24 hours	9	0.50	0.00	1.00	0.35	0.13
Assisted vent with artificial airway	9	0.53	0.00	1.00	0.38	0.15
Active weaning	9	0.94	0.75	1.00	0.11	0.01
Frequent suction: Trach, ETT	9	0.92	0.25	1.00	0.25	0.06
Trach tube	9	0.97	0.75	1.00	0.08	0.01
Chest tube	9	0.94	0.75	1.00	0.11	0.01
Nasal cannula	9	0.72	0.25	1.00	0.32	0.10
Respiratory Isolation	9	0.86	0.50	1.00	0.22	0.05
CARDIOVASCULAR	6	0.81	0.67	0.94	0.09	0.01
Telemetry	9	0.86	0.50	1.00	0.18	0.03
VS > Q Shift	9	0.94	0.75	1.00	0.11	0.01
Fluid challenges	9	0.78	0.50	1.00	0.23	0.05
Continuous IV med administration requiring lab &/or VS monitoring	9	0.83	0.25	1.00	0.25	0.06
IVP anti-arrhythmics	9	0.75	0.25	1.00	0.28	0.08
Initial digitalization	9	0.67	0.00	1.00	0.35	0.13
GASTROINTESTINAL	6	0.81	0.72	1.00	0.11	0.01
Tube feedings	9	1.00	1.00	1.00	0.00	0.00
NGT suction	9	0.72	0.50	1.00	0.23	0.05
Assisted bowel evacuation	9	0.78	0.50	1.00	0.23	0.05
Stools > 1 shift	9	0.78	0.50	1.00	0.26	0.07
Diarrhea	9	0.86	0.50	1.00	0.22	0.05
Colostomy/Ileostomy	9	0.72	0.25	1.00	0.26	0.07
RENAL/GENITOURINARY	5	0.85	0.69	1.00	0.12	0.01
Incontinent	9	0.83	0.50	1.00	0.22	0.05
Standard I & O	9	0.94	0.75	1.00	0.11	0.01
Foley	9	1.00	1.00	1.00	0.00	0.00
Bladder irrigation	9	0.69	0.25	1.00	0.32	0.11
Hemodialysis/Peritoneal dialysis	9	0.81	0.25	1.00	0.27	0.07
SKIN/WOUND CARE	4	0.91	0.83	0.97	0.06	0.00
Complex dressing change	9	0.94	0.75	1.00	0.11	0.01
Decubitus dressing (simple)	9	0.92	0.50	1.00	0.18	0.03
Catastrophic wounds	9	0.83	0.25	1.00	0.25	0.06
Replace soiled dressing	9	0.97	0.75	1.00	0.08	0.01

	N	Mean	Min	Max	Std. Dev.	Var.
NEUROLOGICAL	5	0.79	0.69	0.94	0.11	0.01
Slightly confused	9	0.94	0.75	1.00	0.11	0.01
Confused/disorientated	9	0.86	0.50	1.00	0.18	0.03
Comatose	9	0.78	0.25	1.00	0.32	0.10
Seizure precautions	9	0.69	0.25	1.00	0.30	0.09
Restraints (mitts, vest, wrist)	9	0.69	0.25	1.00	0.27	0.07
MEDICATIONS	4	0.90	0.89	0.92	0.02	0.00
IVP meds/VS documentation	9	0.89	0.75	1.00	0.13	0.02
PRN meds < 4/shift	9	0.89	0.75	1.00	0.13	0.02
PRN meds > 4/shift	9	0.89	0.75	1.00	0.13	0.02
Total # meds adm. 24 hours	9	0.92	0.75	1.00	0.12	0.02
IV'S/LINES	12	0.79	0.61	0.94	0.10	0.01
Insulin drip with Q 1 hour BS	9	0.89	0.25	1.00	0.28	0.08
Central line with TPN	9	0.83	0.25	1.00	0.25	0.06
BS q 6 hours with sliding scale	9	0.89	0.50	1.00	0.18	0.03
BS q 4 hours with sliding scale	9	0.75	0.25	1.00	0.28	0.08
> 2 IV antibiotics	9	0.89	0.25	1.00	0.25	0.06
Routine tubing change	9	0.94	0.75	1.00	0.11	0.01
Continuous sedation drip	9	0.72	0.50	1.00	0.23	0.05
Blood Products this shift	9	0.75	0.25	1.00	0.25	0.06
Blood products previous shift	9	0.72	0.25	1.00	0.29	0.09
Bolus IVPB not scheduled	9	0.69	0.25	1.00	0.27	0.07
Peripheral IV saline lock	9	0.83	0.50	1.00	0.22	0.05
Start peripheral IV	9	0.61	0.50	1.00	0.18	0.03
DAILY ACTIVITIES	17	0.92	0.72	1.00	0.08	0.01
Discharge Teaching	9	0.89	0.50	1.00	0.22	0.05
Total lift to chair	9	0.97	0.75	1.00	0.08	0.01
Total bed bath	9	0.97	0.75	1.00	0.08	0.01
Completely dependent ADLs	9	0.97	0.75	1.00	0.08	0.01
Daily weight	9	0.89	0.50	1.00	0.18	0.03
Psychological Support > 1 hour/24	9	0.94	0.75	1.00	0.11	0.01
Pre-op prep and teaching	9	0.81	0.25	1.00	0.23	0.05
Knowledge deficit/med. teaching	9	0.89	0.50	1.00	0.18	0.03
Admission assessment & history	9	0.97	0.75	1.00	0.08	0.01
Care plan initiation	9	1.00	1.00	1.00	0.00	0.00
Care plan modification	9	1.00	1.00	1.00	0.00	0.00
Feeder	9	0.86	0.50	1.00	0.18	0.03
Linen change (1 x # of changes)	9	0.94	0.75	1.00	0.11	0.01
Frequent turns and position	9	1.00	1.00	1.00	0.00	0.00

	N	Mean	Min	Max	Std. Dev.	Var.
Obesity with immobility	9	0.86	0.50	1.00	0.18	0.03
Cont. Nsg. time 30 to 60 min. (4 x # of nurses)	9	0.89	0.50	1.00	0.18	0.03
Cont. Nsg. time 30 to 60 min. (3 x # of nurses)	9	0.72	0.25	1.00	0.32	0.10
SPECIAL PROCEDURES	6	0.79	0.72	.86	0.06	0.00
Bedside procedures-assist	9	0.86	0.50	1.00	0.18	0.03
Procedures outside hospital	9	0.72	0.25	1.00	0.29	0.09
Code Blue	9	0.81	0.50	1.00	0.21	0.04
Unplanned transfer to IMU/ICU	9	0.78	0.25	1.00	0.29	0.09
Post-op first 24 hours	9	0.72	0.00	1.00	0.38	0.15
MRSA/VRE isolation	9	0.86	0.00	1.00	0.33	0.11
PSYCHOSOCIAL	3	0.94	0.86	1.00	0.07	0.01
Patient call light > 1/hour	9	0.97	0.75	1.00	0.01	0.01
Imminent patient demise	9	0.86	0.25	1.00	0.08	0.08
Frequent family intervention	9	1.00	1.00	1.00	0.00	0.00
TOTALS	78	.8467			1.272	

Major	Minor
PULMONARY	**PULMONARY**
Active weaning	Extubation first 24 hours
Frequent suction: Trach, ETT	Decannulation first 24 hours
Trach tube	Assisted vent with artificial airway
Chest tube	Nasal cannula
Respiratory Isolation	
CARDIOVASCULAR	**CARDIOVASCULAR**
Telemetry	Fluid challenges
VS > Q shift	IVP anti-arrhythmics
Continuous IV med. administration requiring lab &/or VS monitoring	Initial digitalization
GASTROINTESTINAL	**GASTROINTESTINAL**
Tube Feedings	NGT suction
Diarrhea	Assisted bowel evacuation
	Stools > 1/shift
	Colostomy/Ileostomy
RENAL/GENTOURINARY	**RENAL/GENTOURINARY**
Incontinent	Bladder irrigation
Standard I & O	
Foley	
Hemodialysis/Peritoneal Dialysis	
SKIN/WOUND CARE	**SKIN/WOUND CARE**
Complex dressing change	
Decubitus dressing (simple)	
Catastrophic wounds	
Replace soiled dressing	
NEUROLOGICAL	**NEUROLOGICAL**
Slightly confused	Comatose
Confused/disorientated	Seizure Precautions
	Restraints (mitts/vest/wrist)
MEDICATIONS	**MEDICATIONS**
IVP meds/ VS documentation	
PRN meds < 4/shift	
PRN meds > 4/shift	
Total # meds adm./24 hours	
IV'S/LINES	**IV'S/LINES**
Insulin drip with q 1 hour BS	Blood sugars q 4 hours with sliding scale

Central line with TPN	Bolus IVPB not scheduled
Blood sugars q 6 hours with sliding scale	Start peripheral IV
> 2 IV antibiotics	
Routine tubing change	
Continuous sedation drip	
Blood products this shift	
Blood products previous shift	
Peripheral IV saline lock	
DAILY ACTIVITIES	**DAILY ACTIVITIES**
Discharge Teaching	Cont. Nursing time 30 to 60 minutes (3 x # of nurses)
Total lift to chair	
Total bed bath	
Completely dependent ADLs	
Daily weight	
Psychological support > 1hour/24	
Pre-op prep and teaching	
Knowledge deficit/med. teaching	
Admission assessment and history	
Care plan initiation	
Care plan modification	
Feeder	
Linen change (1 x # 1of changes)	
Frequent turns and reposition	
Obesity with immobility	
Cont. Nursing time > 60 minutes (4 x # of nurses)	
SPECIAL PROCEDURES	**SPECIAL PROCEDURES**
Bedside procedures-assist	Procedures outside hospital
Code Blue	Unplanned transfer to ICU/IMU
MRSA/VRE isolation	Post-op first 24 hours
PSYCHOSOCIAL	**PSYCHOSOCIAL**
Patient call light > 1/hour	
Imminent patient demise	
Frequent family intervention	