1994

Development and testing of a patient classification instrument derived from Orem's model.

Pi-Ru Chiu
University of Alabama at Birmingham

Order Number 9520559

Development and testing of a patient classification instrument derived from Orem's model

Chiu, Pi-Ru, D.S.N.

University of Alabama at Birmingham, 1994

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

DEVELOPMENT AND TESTING OF A PATIENT CLASSIFICATION INSTRUMENT DERIVED FROM OREM'S MODEL

by

PI-RU CHIU

A DISSERTATION

Submitted in partial fulfillment of the requirements for
the degree of Doctor of Science in Nursing in the
School of Nursing in the Graduate School,
The University of Alabama
at Birmingham

BIRMINGHAM, ALABAMA

1994

ABSTRACT OF DISSERTATION
GRADUATE SCHOOL, UNIVERSITY OF ALABAMA AT BIRMINGHAM

Degree DSN Major Subject Nursing

Name of Candidate Pi-Ru Chiu

Title Development and Testing of a Patient Classification Instrument Derived from Orem's Model

The purpose of this study was to develop a six-level acuity-based factor-type medical-surgical patient classification system derived from Orem's (1991) model and to test its reliability and validity in an acute-care teaching hospital in the southeast region of the United States. The items and indicators for the instrument were developed from the framework of Orem's model; thus, the instrument was named Orem's Patient Classification System (OPCS). The future goal is to further validate and implement the OPCS in the medical-surgical units of an acute-care teaching hospital in Taiwan.

To test the instrument in the field, a combination of descriptive, factorial, and correlational research design was applied. Three research hypotheses were used to test the instrument for interrater reliability, contrast groups construct validity, and predictive validity.

A 6-patient convenient sample and a 30-patient random sample were selected from five medical-surgical units, one medical intensive care unit, and one surgical intensive care unit. Statistics involved in testing the instrument included percentage agreement, one-tailed z test of kappa statistic, two-way and one-way ANOVA, Sheffé test, and Pearson's correlation coefficient. The tests demonstrated interrater reliability of the overall instrument; however, the interrater reliability of two items of the instrument did not reach a statistically significant level. Also, the tests demonstrated predictive validity and partially demonstrated contrasted groups construct validity of the instrument. The instrument needs further pilot-testing before it is implemented in Taiwan.

Abstract Approved by: Committee Chairman Rachel M. South

Program Director Elizabeth Stullenbarger

Date 12/15/94 Dean of Graduate School Joan L. Lodin

ACKNOWLEDGMENTS

My gratitude shall always be extended to Dr. Rachel Booth, Dean of the School of Nursing and chair of my dissertation committee, for enthusiastically encouraging, mentoring, and motivating me throughout my research process and for spending countless hours advising and refining my work. Without her splendid support, this work would never have been completed.

I would like to express my special appreciation to members of my dissertation committee. Dr. Linda Davis, Assistant Dean for Undergraduate Studies in the School of Nursing, directed in my research methodology and edited my work to improve the clarity of my study. Dr. Elizabeth Stullenbarger, Associate Dean for Graduate Studies in the School of Nursing, guided me with her knowledge related to Orem's model and provided me with great emotional support. Dr. Sean-Shong Hwang shared his expertise in statistics with me and always had a positive attitude. Dr. Howard Houser willingly spent hours reviewing my manuscript and provided valuable suggestions. Dr. Janice Gay, former member of my dissertation committee, assisted me in editing and organizing my work.

My special appreciation goes to other faculty members and researchers of the school. Dr. Carol Dashiff and Dr. Nena Sanders served as content experts for this study. Dr. Jiang-Nan Wu assisted the study by writing a statistical formula into a computerized program for statistical analysis. Dr. Malcolm Turner guided the study by sharing his wealth of knowledge in statistical analysis.

I wish to express my gratitude to the staff at the University of Alabama at Birmingham Hospital for their support of research in the field of medical-surgical units and especially to the director of the nursing department and the director of the hospital's nursing research committee for being so supportive throughout the study.

I wish to express my appreciation to Barbara Morales and Dories Godsey for their participation in the data collection procedures. I also wish to express my appreciation to Stephanie Burnett, Head Nurse at Spain Rehabilitation Center, for her administrative support. All participants in this study are thanked for their willingness to make their contributions to this study.

Finally, I deeply appreciate my friends, my colleagues, and my parents, who are always support me through Graduate School.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

LIST OF TABLES

LIST OF TABLES (Continued)

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Patient classification is a process of grouping patients into a number of categories according to some measure of patient property, characteristics of units, or nursing care requirements (Dijkers & Paradise, 1986; Giovannetti, 1979; Polit & Hungler, 1991). Patient classification systems (PCSs) can be defined as the instruments designed to classify patients into groups based on patient care requirements to assist in making decisions about nurse staffing, patient assignment, case mix analysis, nursing cost per patient, budget planning, variable billing, and maintenance of nursing quality (De Groot, 1989a).

The use of patient classification systems permits measurement and quantification of the acuity of patient illness and level of nursing care required (Sullivan & Decker, 1985). Data obtained from these systems are useful, because qualified registered nurses, licensed practical nurses, or nurse aides can be assigned according to the acuity levels of patients and patients' care needs (Lewis & Carini, 1984). Most importantly, patient classification systems provide one of the most widely used methods of measuring workload for the determination of appropriate allocation of

nurse staffing (Dijkers & Paradise, 1986; Giovannetti, 1979; Hastings, 1987; Joint Commission on Accreditation of Healthcare Organizations [JCAHO], 1982; Nagaprasanna, 1988), especially in acute care hospitals where patient variables are the most important determinants of nursing workload (Giovannetti, 1979).

In the last few decades, a number of systems have been developed for different care providers and for various purposes (McKenzie, 1991). However, no single instrument is considered to be completely satisfactory in fulfilling its mission (De Groot, 1989a; Nagaprasanna, 1988).

The purpose of this study was to develop a six-level (type VI) acuity-based factor-type patient classification system for medical-surgical patients between 18 and 65 years of age using a theoretical framework derived from Orem's (1991) self-care model and to pilot test the instrument for reliability and validity in an acute-care teaching hospital in the southeastern region of the United States. To make the data manageable, only the adult medical-surgical patients were selected from the population during the initial stage of instrument development. The age range from 18 to 65 indicated that patients selected were in their adulthood (Erikson, 1963). The future goal is to test and implement this newly developed patient classification system in the medical-surgical units of an acute-care teaching hospital in Taiwan.

Background

The idea of patient classification can be traced back to 1863, when Florence Nightingale identified the value of differentiating between medical and surgical patients because of their unique and separate needs (Davis & Bertram, 1991; Jackson & Resnick, 1982). In 1947, the National League for Nursing Education published a factor-type patient classification system that used a three-point scale for classifying pediatric patients' needs (Alward, 1983). At that time, estimation of nursing time required for patients in each class was constructed in gross terms only and, thus, could not provide a sensitive tool for the determination of staffing needs (Abdellah & Levine, 1979). During the early 1960s, the historical staffing patterns based upon census count or trial-and-error were criticized by industrial engineers as no longer properly reflecting the staffing needs of patients (Des Ormeaux, 1977; Kessler, Kessler, & Knibloe, 1990; Mowry & Korpman, 1986). Thereafter, the first generation of a patient classification system was developed in the 1960s at the Johns Hopkins University to predict nurse staffing levels needed for each shift (Giovannetti & Johnson, 1990; Wolfe & Young, 1965). The use of PCSs became widely accepted as a means to determine appropriate nursing staffing patterns throughout the 1970s (Roehrl, 1979).

For 20 years after the first PCS was developed, the primary purpose of the hospital inpatient classification system was to determine nurse

staffing (Nagaprasanna, 1988). Thereafter, the functions have expanded from nurse staffing to include patient assignment, case mix analysis, nursing cost per patient, budget planning, variable patient billing, and maintenance of nursing quality (Budd & Propotnik, 1989; De Groot, 1989a; Hoffman, 1988; Slyck, 1991; Stepura & Miller, 1989; Swansburg & Sowell, 1992). In general, the implementation of patient classification systems leads to a better distribution of staff nurses, maintains nursing quality, and contributes to more cost-effective functioning of the nursing department (Kyle & Kinder, 1990).

Interest in PCSs increased in the 1980s following the development of a nursing standard by the Joint Commission on Accreditation of Health Organizations (JCAHO, 1982) that addressed this issue. In 1982, the JCAHO indicated that the nursing department should define, implement, and maintain a system for determining patient requirements for nursing care. In 1992, the statement was further expanded to require that the hospital-wide patient care programs, policies, and procedures should describe how the nursing care needs of patients were assessed, evaluated, and met (JCAHO, 1992).

The demand from JCAHO (1992) for better methods to identify staffing needs precipitated the development of PCSs to consider specialty areas of nursing care (Davis & Bertram, 1991; Niemeier & Reed, 1985). The development of indicators provided a mechanism for designing a

system that would address these specialty needs. An indicator depicts "separate but related dimensions of patient care which, when considered together, predict the level of patient care likely to be required" (De Groot, 1989a, p. 31). Indicators used in a patient classification system must be simple and efficient or the system will become too complex and time-consuming to complete accurately (De Groot, 1989b). Even when the system is simple and efficient, to have one system apply to all units within a hospital engenders another problem, that is, the inability to capture the distinctive characteristics and intensities of specialized care needed in different nursing units (Cottey, Nauert, & Willis, 1992). Therefore, systems have been developed to meet the unique needs of specialty units (O'Leary, 1991; Schwamb, 1989). Some examples include: (a) medical-surgical (Strom, 1993; Van Der Walt, 1992), (b) critical care (Niemeier & Reed, 1985), (c) burn intensive care (Cottey et al., 1992), (d) oncology (Arenth, 1985; Lovett, Wagner, & McMillan, 1991), (e) dialysis (Kessler et al., 1990), (f) maternity (Schwamb, 1989), (g) psychiatric (Croft, 1993; O'Leary, 1991; Schroder, Washington, Deering, & Coyne, 1986), (h) rehabilitation (Gender, 1989), and (i) ambulatory (Hastings, 1987; Hoffman & Wakefield, 1986; Johnson, 1989). As a result, there may be as many as five different systems being used simultaneously in a given hospital (Slyck, 1991).

Mean care hours required per specialty area varies among units and/or institutions due to differences in patient needs and the specific

circumstances of each unit/institution. For example, a pediatric type IV system showed the mean nursing hours needed per 24 hours from Classes I through IV as 1.86, 3.22, 5.45, and 7.53, respectively (Chagnon, Audette, Lebrun, & Tilquin, 1978), whereas a medical-surgical type VI system showed the mean nursing hours needed from Classes I through VI as 1.00, 3.01, 4.6, 6.81, 10.65, and 12.8, respectively (Van Der Walt, 1992), and an oncology type V system showed the mean nursing hours from Classes I through V as high as 4.0, 6.0, 10.0, 14.0, and 24.0, respectively (Arenth, 1985).

Levels used to classify patients were also expanded from three in 1965 (Wolfe & Young, 1965), to four in the 1970s (Roehrl, 1979), and to five in 1985 (Arenth, 1985). In 1988, Nagaprasanna conducted a survey of 213 hospitals which were randomly sampled from 700 hospitals with a bed capacity of more than 400 and found that 22% of the hospitals used three levels (three-patient category), 55% of the hospitals used four levels (four-patient category), and 22% used five levels (five-patient category). With increasing technology and higher acuity of patients (Van Der Walt, 1992), it was believed that five levels were still too limited to reflect the variability of care intensity that patients required (Dijkers & Paradise, 1986). Thus, the newest system uses six levels to classify patients (Croft, 1993; Johnson, 1989; Van Der Walt, 1992). The acuity level of patients continues to increase in hospitals, as reported by Gender (1989) and Batty, Mooney, and

Lowry (1990). The acuity level of patients is "the relative amount of nursing care required as determined by a therapeutic indicator within a level of care" (Arenth, 1985, p. 18). Gender (1989) found that the acuity level of hospitalized patients increased from an average of 2.7 in 1987 to 2.9 in 1988, whereas Batty et al. (1990) found an increase from an average of 1.75 to 2.00 in a 1-year period in 1989. The changes in patients' acuity levels and the concurrent increase in levels of systems demonstrate that these factors have been the stimulus for frequent modifications and/or development of new patient classification systems.

Patient classification is at the heart of any staffing workload analysis system (Vaughn & MacLeod, 1980). The dramatic changes in patient acuity, nursing requirements, and health care practices make it imperative that PCSs be evaluated annually and modified as needed to capture realistic nursing workload (Cottey, Nauert, & Wills, 1992).

The exact number of patient classification systems currently in use is not known. In 1973, Aydelotte reported 40 types of systems. In 1979, Giovannetti estimated that 1,000 hospitals were using some form of patient classification system. In Nagaprasanna's (1988) report, 34 (16%) of the hospitals used commercialized patient classification systems such as the Commission for Administration Services in Hospitals (CASH) and Medicus. In 1990, Wake reported on a survey of 987 hospitals in the United States and showed that over 100 commercially available systems were in use. The

most frequently used systems were Grace-Reynolds Application and Study of PETO (GRASP) (17%) and Medicus (11%).

Pardue and Dick (1986) reported that approximately 80% of hospitals developed their own systems that were based on systems in use elsewhere. In Nagaprasanna's (1988) report, 40% of patient classification systems were internally developed. However, Wake's (1990) showed that in 1986, only 32% of the hospitals reported using internally developed systems, and in 1989, only 31%.

In 1981, Huckabay and Skonieczny published a survey of 236 hospitals in which 28 (12%) reported that they had not implemented a patient classification system. In 1990, Wake's report showed that 13% of the hospitals reported having no system in 1986, 4% having no system in 1989, and 1% projected they would acquire one by 1992.

Statement of the Problem

Although numerous systems have been developed, implemented, and modified in hospitals over the past few decades (Aydelotte, 1973; Giovannetti, 1979; Wake, 1990), only 62% of the hospitals were satisfied with the systems they were using (Nagaprasanna, 1988). No single instrument is considered to be completely satisfactory for classifying patient needs and for guiding the nurse administrator in making decisions for nurse staffing (De Groot, 1989a; Nagaprasanna, 1988; Zembala, 1993).

The need to develop further and to refine patient classification systems persists because (a) the domain of nursing is inherently complicated and is difficult to measure (King, 1975), (b) most systems currently in use are based on various modifications of early industrial engineering methodology (Jennings, Rea, Antopol, & Carty, 1989), and (c) only two reports have been identified that applied nursing models to the development of PCSs (Donnelly, 1981; Leatt, Bay, & Stinson, 1981). Because of these circumstances, nurses who are developing PCSs must strive to generate more nursing-domain data and to explore appropriate systems for nursing staffing purposes.

Significance of the Study

The importance of basing the development of a patient classification instrument on nursing theory has been recognized by some nurse leaders (Giovannetti, 1978; Haas, 1988). Giovannetti (1978) proposed that patient classification systems should be based on a theoretical framework that reflects the nursing perspective. Haas (1988) indicated that

> the construction of a patient classification tool should be grounded in a theory or theories which explain the work of nursing. Only then will the items comprising the patient classification instrument adequately survey the domain of nursing work. (p. 61)

The use of Orem's (1991) self-care model for identifying patient needs provides the structure, definition, and rationale for the development of a patient classification instrument that addresses the complex spectrum

of nursing practice. Consequently, Orem's model was deemed important and practical and was used as the theoretical framework for the development of the PCS in this study. Additionally, a valid and reliable patient classification system that is tailored to a specific population is not available for use in the teaching hospitals in Taiwan. Therefore, due to the variations in patient needs, increase in acuity levels, and differences in nursing requirements, it is imperative that a system be developed.

Statement of Purpose and Research Question

The purpose of this study was to develop a six-level acuity-based factor-type PCS for medical-surgical patients between 18 and 65 years of age, using a theoretical framework derived from Orem's (1991) self-care model and to pilot-test the instrument's reliability and validity in an acute-care teaching hospital in the southeastern region of the United States. The future goal is to further pilot test and implement this newly developed acuity-based system in the medical-surgical units of an acute-care teaching hospital in Taiwan. Therefore, the research question was, what is the reliability and validity of a patient classification instrument that is derived from Orem's self-care model for use in an acute care teaching hospital?

Orem's Self-Care Theoretical Framework

Orem's self-care model provided the theoretical framework for the development of a patient classification instrument for this study. This

section reviews Orem's model and the framework that emanated from the model.

Orem (1991) viewed nursing as a means of assisting individuals in the maintenance of self-care. Her model of the self-care deficit of nursing is comprised of six core concepts: self-care, self-care agency, therapeutic self-care demand, self-care deficit, nursing system, and nursing agency, and one peripheral concept, basic conditioning factors. The model is expressed in three related parts or theories that include self-care, self-care deficit, and nursing system (Orem, 1985; 1991).

The three theories are closely related, with the theory of nursing system encompassing the theory of self-care deficit, and the theory of self-care deficit encompassing the theory of self-care. For example, the concept of nursing system will have no meaning in the absence of the concepts of self-care and self-care deficit. Thus, Orem's model can be described in sequences from theory of self-care, theory of self-care deficit, to theory of nursing system, and is depicted by the six core concepts and one peripheral concept, which are linked among the three interrelated nursing theories.

Self-Care

The theory of self-care indicates that individuals need to act on their own behalf to meet the requisites for maintaining life and functioning. Concepts included in the theory of self-care are self-care, self-care agency,

therapeutic self-care demand, basic conditioning factors, and self-care requisites.

Self-care is a learned behavior, a form of deliberate action, and occurs as individuals engage in actions to take care of themselves. Self-care agency is the complex, developed capability that enables individuals to regulate their functioning and development and to perform the care measures required for self-care. Therapeutic self-care demand are measures of care required to maintain life and promote health.

An individual's self-care agency is affected by basic conditioning factors, elements that affect an individual's ability to engage in required self-care. Orem (1991) identified 10 basic conditioning factors: (a) age, (b) gender, (c) developmental state, (d) health state, (e) sociocultural orientation, (f) health care system, (g) family system, (h) pattern of living, (i) environment, and (j) available resources.

Self-care requisites form the basis for initiating self-care; they are the intended or desired results of self-care. Self-care requisites can be classified into three categories: universal, developmental, and health-deviation. Originally, development and health-deviation self-care requisities were considered to be constituents of universal self-care requisites. They have been segregated to highlight their importance and characteristics (Orem, 1991).

Universal self-care requisites represent the kind of actions needed by all persons to maintain life and regulate functioning, such as maintenance of sufficient intake of air, water, and food; provision of basic care associated with the elimination processes and excrements; and maintenance of a balance between rest and activity. Developmental self-care requisites are associated with the desire for conditions that support human development or the needed help for conditions that adversely affect human development during the human developmental cycle, such as childhood, pregnancy, terminal illness, and death of a family member. Health-deviation self-care requisites discern needed actions that individuals long for when they are ill or injured and are under medical treatment. For example, the individual needs to seek and secure appropriate medical assistance; to be aware of and attend to the effects of pathologic conditions; and to carry out prescribed diagnostic, therapeutic, and rehabilitative measures.

Self-Care Deficit

The theory of self-care deficit characterizes individuals who are unable to perform the actions of self-care. This theory depicts the relationship between the concept of self-care agency and concept of therapeutic self-care demands.

The relationship between individuals' self-care agency in performing self-care and their therapeutic self-care demands can be explained in terms of equal to, less than, and more than. When therapeutic self-care demand

is greater than self-care agency, the individual is in a state of self-care deficit and needs assistance from others.

Nursing System

The nursing system is an action system that is constructed through actions of nurses and patients. If individuals have self-care deficits and their needs cannot be met by family members, nurses who are operating their nursing agency under nursing systems can assist the individuals in meeting their therapeutic self-care demands and in restoring or developing their capacities for self-care. Nursing agency, a complex property of nurses, is exhibited through the deliberate estimations of both patients' self-care deficits and basic conditioning factors, and by applying helping methods such as providing care, teaching, guiding or supporting the individual, or modifying the environment.

Nursing systems can be either wholly or partly compensatory, or supportive-educative. When individuals are unable to care for themselves, they are placed in a wholly compensatory system, in which nurses devote most of their helping methods for the individuals. When individuals can act partially for their own self-care, they are placed in a partly compensatory system, wherein nurses help them carry out self-care in areas of deficit. As individuals assume most of the actions of self-care, they are moved to the supportive-educative system, wherein the individuals perform most of their self-care actions and nurses help them restore or develop their self-care

agency through guidance, education, and support. Thus, the theory of nursing system signifies a nursing agency in which individuals' self-care deficits are identified and their therapeutic self-care demands are met.

A valid PCS should have the power to identify patients' self-care deficits and provide a system in which nurse staffing can be determined in a cost-effective manner. In reviewing Orem's model, two frameworks were identified and selected as guidelines for the construction of the PCS:

1. use of the self-care requisites for development of the instrument items, and

2. use of the three major concepts of self-care agency, self-care deficit, and nursing agency for formulation of the indicators.

The instrument derived from Orem's self-care nursing model has the ability to identify and measure patients' self-care deficits and to calculate nurses' workloads in meeting patients' therapeutic self-care demands; therefore, it can be a useful instrument for the determination of appropriate nurse staffing.

To recognize Orem's contribution to this system, this newly developed instrument is called the Orem Patient Classification System (OPCS). A schematic illustration (see Figure 1) indicates the theoretical framework used in the PCS derived from Orem's model, and the dynamics and the relationships among self-care deficits, self-care agency, therapeutic

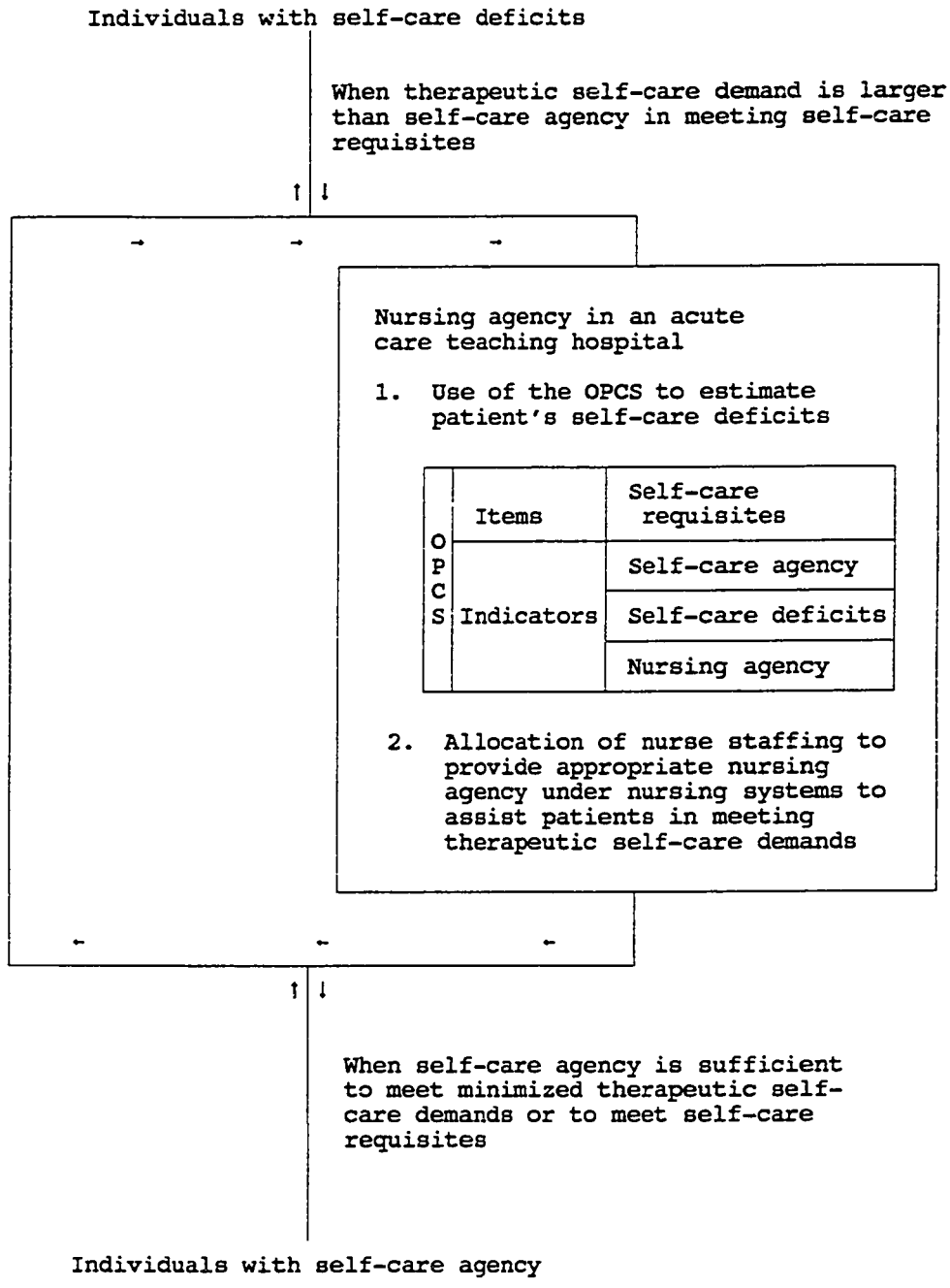


Figure 1. Theoretical framework used in OPCS derived from Orem's model in an acute care teaching hospital. Note the dynamics and relationships among self-care agency, self-care requisites, therapeutic self-care demand, self-care deficits, nursing agency, and nursing systems.

self-care demand, and nursing agency. The importance of the OPCS is also highlighted as a necessary procedure before appropriate nursing actions can be implemented.

Steps Taken in Researcher-Developed Instrument

The purpose of this study was to develop and test a six-level patient classification instrument derived from Orem's (1991) self-care model. An acuity-based factor-type was selected as the method of measurement. The instrument was developed to determine appropriate nurse staffing based upon the amount of care needed by patients through the identification of patients' level of self-care requisites or deficits. The detailed processes and methodology related to the development of the instrument are described more fully in Chapter III; however, the nine steps required for instrument development and testing are described briefly below.

1. Information relevant to the development and testing of the PCS was reviewed and included a review of PCSs, nursing theories, analysis of existing instruments, and consultation of experts in instrumentation.

2. First draft of the instrument was developed according to the following four steps:

   a. Terms from Orem's model were defined.

   b. Items from Orem's model were constructed and revised.

   c. Indicators, categories, and a scoring system for each item were developed.

d. Directions and guidelines for the use of the instrument and patient classification form (PCF) were developed.

3. Instrument was tested for content validity by two content experts using the Content Validity Index (CVI) method (Waltz, Strickland, & Lenz, 1991).

4. Instrument was revised according to the feedback from content experts.

5. Instrument was retested for content validity following previous procedure.

6. Training sessions for registered nurse raters were conducted.

7. Instrument was pre-tested for interrater reliability using percentage agreement estimation with a small sample of six hospitalized medical-surgical patients.

8. Instrument was tested again in the field (pilot-testing) on a sample of 30 hospitalized medical-surgical patients. Tests used for the field study included interrater reliability, contrast groups construct validity, and predictive validity.

9. A six-level acuity-based factor-type medical-surgical Orem Patient Classification System (OPCS) was established according to the proportion of required nursing time estimated for each classification level. Data collected for establishing this system are from the same sample used in the pilot testing of the instrument.

Assumptions

The assumptions on which this study was based are as follows:

1. Registered nurses are able to identify patients' self-care deficits and judge the extent of their nursing needs.

2. Registered nurses are able to estimate the number of patients they can care for based upon the information about the patients' nursing care needs.

3. The nursing hours required for patients in the medical-surgical intensive care units are greater than those for patients who are not ready for discharge from the medical-surgical units within 48 hours.

4. The nursing hours required for patients who are not ready for discharge from the medical-surgical units within 48 hours are greater than those for patients who are ready for discharge from the medical-surgical units within 48 hours.

5. Testing of the researcher-developed instrument could serve as the basis for future testing in a hospital in Taiwan.

Research Hypotheses

The purpose of this study was to develop a six-level acuity-based factor-type PCS for medical-surgical patients who were between 18 and 65 years of age using a theoretical framework derived from Orem's self-care model and to test the reliability and validity of the OPCS. The reliability of the instrument was tested using interrater reliability. The validity of the

instrument was tested using content validity, contrasted groups validity, and predictive validity. Content validity was first determined by using content experts. After the final content of the instrument satisfied the preset criterion, other methods of testing were conducted. Thus, the research hypotheses utilized to test the instrument included: interrater reliability, contrasted groups validity, and predictive validity.

A research hypothesis is a statement of relationships among variables that the researcher predicts (Mateo & Kirchhoff, 1991). This study applied three research hypotheses; the first tested the instrument for interrater reliability; the second, for contrasted groups construct validity; and the third, for predictive validity. The hypotheses are stated below.

Hypothesis 1: There will be a significant interrater agreement on items of the OPCS among the nurse raters (interrater reliability).

Hypothesis 2: Patient total classification scores of the OPCS will demonstrate a high-to-low sequence among patients in medical-surgical intensive care units, patients not ready for discharge from medical-surgical units within 48 hours, and patients ready for discharge from medical-surgical units within 48 hours (contrasted groups construct validity).

Hypothesis 3: Using the OPCS, there will be a positive correlation between patient total classification scores and the nursing care time required by patients (predictive validity).

Summary

Patient classification systems have been recognized as useful and important instruments in the determination of appropriate allocation of nurse staffing in hospitals. However, no one system is entirely satisfactory, and little research has been reported that indicates the use of a nursing model or theory in the development of a PCS. Orem's self-care model serves as a guide for nurses in assessing patient self-care agency, identifying patient self-care deficits, and assisting patients to retain their self-care ability. Therefore, Orem's (1991) self-care model was used as a theoretical framework in the development of a patient classification instrument that is believed by this author to reflect the complex domain of nursing practice and to provide support for the use of nursing theory in practice.

In this chapter, background of patient classification systems has been presented, followed by a description of the problems currently encountered, significance of the study, purpose of the research, and research question. The final part of the chapter delineates the theoretical framework of Orem's (1991) model, steps taken in the instrument development, and assumptions and hypotheses of the research.

CHAPTER II

LITERATURE REVIEW

This chapter presents the findings from efforts made to develop patient classification systems, frameworks used for system development, and methods in the testing of the system. Important considerations of user training, influence of computer, and comparisons among systems and timing of classification are addressed. A summary for the chapter is provided.

Development of Patient Classification System (PCS)

There are many different sources of patient classification systems used in hospitals. Some are designed commercially, while others are either developed internally by hospital staff or are modifications of systems in use elsewhere (Nagaprasanna, 1988; Pardue & Dick, 1986; Wake, 1990). To develop a patient classification system for nurse staffing purposes, four interrelated aspects need to be considered: methods of indicator measurement, types of patient classification systems, levels of patient classification, and characteristics of hospital units. These considerations are described in the following sections.

Methods of Indicator Measurement

An indicator is developed to "depict separate but related dimensions of patient care which, when considered together, predict the level of patient care likely to be required" (De Groot, 1989a, p. 31). With different types of indicators developed for each system, measures on nursing time required per patient class may vary from system to system. Indicators used in a system can be categorized as disease-based, acuity-based, procedure-based, or a combination of parameters of the three (Bermas & Slyck, 1984). These categories are described below.

Disease-Based

In systems using disease-based indicators, patients are divided primarily according to the types of diseases. Disease-based systems are medically focused rather than nursing focused, and patients are divided according to types of diseases. These systems are used to fulfill medical or reimbursement purposes; they are not intended to determine day-to-day nurse staffing. Examples of the disease-based systems include the International Classification of Diseases, 9th Revision, Clinical Modifications (ICD-9-CM List A) (Hoffman, 1988), and Diagnosis Related Group (DRG) (Fetter, Brand, & Gamache, 1991; Fetter, Mills, Riedel, & Thompson, 1977; Rosko, 1988). Both of these systems were designed to group patients into classes using similar amounts of resources.

ICD-9-CM List A includes a list of 398 diagnostic groups, which are further subcategorized into a total of 7960 case types (Hoffman, 1988). With thousands of categories being used, this system provides fairly exact classification. However, it is less helpful in doing statistical analyses of patients for various activities. This type of classification is widely used for reimbursement purposes.

Diagnosis related groups (DRG) provides a list of 468 diagnostic groups, a more manageable data set than the ICD-9-CM List A (Rosko, 1988). The amount and type of nursing required by patients in certain DRGs has been found to be variable (Mowry & Korpman, 1985). For example, nursing labor costs per day can vary from $42 to $223 for DRG 294 (diabetes), and from $41 to $179 for DRG 182 (gastrointestinal disorders). This type of system cannot accurately reflect the nursing workload required by patients in all groups (Jelinek, 1989). Thus, this category may be suitable for medical or reimbursement purposes, but it is not best for planning day-to-day nurse staffing.

Acuity-Based

The acuity-based system is used to classify patients according to the extent of self-care or the level of assistance needed from the nurse (Huckabay, 1981). Some of the indicators used in the system are patient's physical restrictions, degree of self-sufficiency or dependence, instructional needs, psychosocial and emotional state, and required nursing procedures.

Scores assigned to indicators are weighted to reflect the intensity or hours of nursing care required by patients (Pardue & Dick, 1986), or each indicator within each item is assigned a different weighted score to represent the intensity or hours of nursing care. The primary purpose of these systems is to predict nurse staffing needs; they are not designed to determine nursing care costs.

Factors used to measure patient acuity level may include activities such as "patient's capacities in carrying out the activities of daily living, need for treatments and medications and emotional/behavioral state" (Dijikers & Paradise, 1986, p. 25). By this definition, it is clear that an acuity-based system may include indicators that identify either the patient's capacities for self-care or nursing procedures such as treatments and medications. However, in the acuity-based system, the emphasis is on the patient's needs for receiving care, treatments, and medications.

Acuity-based systems are used to measure nursing care actually needed by patients and patient acuity level (Sullivan & Decker, 1985). The acuity level of patients is "the relative amount of nursing care required as determined by a therapeutic indicator within a level of care" (Arenth, 1985, p. 18). For example, some systems define the acuity level of Class II patients as 1.0, whereas, the acuity levels for patients in other classes of the same system is based upon comparisons of the amount of nursing care required relative to that of Class II patients (Medicus Systems Corporation,

1990). The advantage to this approach is that, when the standard nursing care time required by the Class II patients of any specific unit is established, the nursing care time required for patients in other classes of the same unit can be easily predicted based upon the pre-established acuity ratio among the patient classes.

The other advantage of acuity-based systems is that the system determines the amount of nursing care a patient should have rather than the care a patient has received (Dijkers, Paradise, & Maxwell, 1986). Thus, these systems seem to be more accurate methods for assessing patients' needs and are considered to be better designed for staffing purposes.

However, several disadvantages of the acuity-based systems have been identified. Because these systems classify patients based upon nursing care needed, not the nursing care actually received, costs could be determined from an acuity-based system only if the assumption were made that the care required is equal to the care received, or to conduct an additional study of nursing care activities in conjunction with the acuity-based system (Sherman, 1990). In other words, if the nurse administrator wants to use a PCS as a basis for determining nursing costs, the instrument must be designed for that specific purpose. An acuity system that reflects exact nursing care hours does not exist (Gender, 1989).

To determine nursing cost for the system, several different methodologies have been used to determine the amount of nursing time

needed per patient classification. Thompson and Diers (1991) suggested a method of gaining consensus on the part of the nursing staff in setting time standards. Sherman (1990) suggested that patient classification data be used in conjunction with data from work sampling to estimate costs (Sherman, 1990). Alward (1983) applied two basic approaches to work analysis of time requirement for each class of patients: time motion observation study and self-recording of activities.

Other criticisms are aimed at the lack of comprehensiveness and a systematic approach in using the critical indicators to assess patient needs (Giovannetti, 1979). Reports of most studies published in the literature do not provide sufficient information about the development of the indicators, instrument, and other essential elements of the system (Leatt et al., 1981).

Procedure-Based

Procedure-based systems are used to measure nursing workload primarily according to nursing tasks rendered, using indicators such as diet, toileting, bathing, vital signs, turning, medications, and suctioning (Meyer, 1978). The time needed for each procedure or task is established by expert judgement or time and motion studies on data obtained from self reports or direct observations (Dijkers & Paradise, 1986; Meyer, 1978). Nursing time is determined for both direct and indirect nursing tasks, frequency of nursing tasks, and other factors, such as fatigue, that consume nursing time during the shift.

The advantage of a procedure-based system is its utility in a cost analysis. It is the indicator that represents the amount of time required to perform the procedure or task; it could be useful in calculating the cost of nursing care (Dijikers & Paradise, 1986).

A disadvantage is the fact that procedure-based systems cannot determine nurse staffing needs, especially when the unit is already understaffed. When the unit is understaffed, which is common in hospitals, classifying patients according to "the care the patient likely received perpetuates understaffing of the unit because it affirms the status quo" (Dijkers & Paradise, 1986, p. 28). Therefore, if this type of system is to be used for staffing purposes, a study should be conducted while the unit is adequately staffed (Williams, 1988). The discrepancy between the hours of care patients need and the hours of care patients actually receive limits the application of these systems in projecting the staffing needs of health care institutions. Another disadvantage of this system is that it reduces the practice of nursing to a series of tasks. In fact, time is not the most important variable for valuing the nursing services provided (Slyck, 1991).

Combined Parameters-Based

Any combination of three approaches may be used in constructing a patient classification system. An example of such a combination is the Acute Physiology and Chronic Health Evaluation II (APACHE II) classification system (Wagner & Draper, 1984). The APACHE includes

indicators that measure patient health status, derangement of physiologic systems, and therapeutic interventions rendered. For example, patient's age, physiological parameters, Glasgow coma scale, nature of any surgery performed, and patient's diagnosis are used to classify the patient (Marks, Simons, Blizzard, & Browne, 1991). However, this system is used to determine severity of illness for predicting morbidity, mortality, and incurred cost; thus, it cannot be applied to determine day-to-day nurse staffing (Ambutas, 1987).

In conclusion, systems with disease-based and combined parameters-based indicators are not used for nurse staffing purpose. Systems with either acuity-based or procedure-based indicators are often used for calculating patient nursing needs or nurses' workloads. Systems with acuity-based indicators are more accurate for assessing patients' needs than those of procedure-based, but the use of a systematic approach in the development of the acuity-based indicators cannot be neglected. Procedure-based systems can measure precisely the amount of nursing time spent in providing care. However, it is critical that the unit be adequately staffed when measuring nursing time needed for each procedure-based indicator.

Types of Patient Classification Systems

The second aspect one needs to consider in the development of an instrument is the decision as to which type of patient classification system

to select. Abdellah and Levine (1979) recognized two types: the prototype and the factor evaluation. Lewis and Carini (1984) also recognized three types of systems: descriptive style, checklist style, and time or relative value unit style. All of these types are described with comparisons and contrasts made when possible.

Prototype

The use of prototype systems can be traced back to the early 1950s in Harper Hospital in Detroit (Wright, 1954) and Walter Reed Army Hospital in Washington, DC (Claussen, 1955). A prototype system is composed of several levels of ranked, mutually exclusive, and exhaustive categories with general descriptions of patients in each category (Mowry & Korpman, 1986). Usually these classes are arranged in an ascending scale in accord with the elevation of nursing workload (Hoffman, 1988). Each class is fully narrated in terms of the characteristics of the typical patient in that category (Abdellah & Levine, 1979). These narrated indicators may be acuity-based, procedure-based, or a blend of both. The patient is classified by assignment to the class that best matches patient characteristics or nursing needs with the indicators described in that class. The nursing hours needed per patient class is determined by additional studies, such as time-and-motion, to obtain the mean nursing hours required by patients under each classification.

Abdellah and Levine's prototype and Lewis and Carini's descriptive type possess essentially the same characteristics. In the latter type, a concise and narrative description is made about the patient for each of the categories and the patient is matched with the particular category that most closely describes the level or intensity of care that the patient needs or receives.

A maternity patient classification system (MPCS) used in Mid-Island hospital is an example of the prototype approach (Schwamb, 1989). MPCS was modified from a prototype system previously used throughout the same hospital. This prototype system contained three levels: Class I included indicators with simple, routine nursing interventions; Class II included indicators such as routine care of IVs, Levine tube, and isolation procedures; and Class III included indicators such as preparation for surgery, frequent monitoring of IVs, vital signs, and vaginal bleeding. After the MPCS was implemented in the maternity unit, it was found that the use of the new system was more cost-effective than was the former system to determine nurse staffing. However, there was not enough data presented in the project to support the reliability and validity of the new instrument.

Although somewhat aberrant from the prototype approach, an Allocation, Resource Identification and Costing (ARIC) system was recently developed (Giovannetti & Johnson, 1990; Johnson, 1989). In this system, nursing workloads were determined by combining results from two

modules of the instrument: dependent needs and independent needs. Dependent needs were those indicators under the prototype approach and were driven by doctors' orders or hospital nursing department policies, such as giving medications and taking blood pressure. Independent needs were indicators that identified the patient's or family's needs and were ranked from levels of very low to extremely high for education, psychosocial support, and coordination of care. A combination of these two components provided various results for determining the appropriate class for each patient.

The value of the prototype approach is its simplicity. However, this system has limitations in that classifying patients based upon a gross assessment of patient characteristics threatens the instrument's reliability and validity. Further, insufficient descriptions of patient needs in each category may result in a biased assessment and inadequate precision for costing purposes (Ruman & Nelson, 1987). As a result, there is a trend in hospitals to convert from prototype to factor evaluation systems (Abdellah & Levine, 1979; Helmer & Halbert, 1987; Pardue & Dick, 1986).

Factor Evaluation

The factor evaluation approach of patient classification classifies patients based upon a total score that is summed from scores ratings on a group of indicators and factors. Indicators used may be acuity-based, procedure-based, disease-based, or combined parameters-based (Bermas &

Slyck, 1984). However, only the acuity-based and the procedure-based indicators are relevant to nursing staffing and, thus, are treated as subsystems for factor evaluation (Abdellah & Levine, 1979). In addition, these two subsystems also correspond with the checklist style and the relative value unit (RVU) style proposed by Lewis and Carini (1984). The following section examines these two subtypes for the purpose of comparing and contrasting them with those proposed by Lewis and Carini.

Acuity-based factor evaluation is the first subtype of factor evaluation. This system classifies patients using acuity-based indicators that have predetermined ranked or weighted scores. Patients are assigned to a class based on total scores obtained from the selected indicators.

For ease of data management, all possible total scores are further grouped into several classes to reflect different nursing care time required within each class (Des Ormeaux, 1977). For example, Class I (minimal care) = 6 - 9 points, Class II (standard care) = 10 - 13 points, Class III (intensive care) = 14 - 17 points, and so on (Abdellah & Levine, 1979). Thus, the total patient score determines the patient's category.

The characteristics of this subtype are analogous to those of Lewis and Carini's (1984) checklist style, which is a checklist-style acuity-based subtype of factor evaluation. This system requires a judgment to be made about the patient's condition and needs using the acuity-based indicators within each item. The scores obtained for the indicators are summed and

matched with the class in which the patient will be placed. Examples of this system are described in the following section.

Van Der Walt (1992) depicted a six-level system for use in a surgical pavilion with 731 beds and a medical pavilion with 501 beds. Patients' requirements for nursing care were assessed using a group of acuity-based indicators which were arranged under six items: (a) assistance with physical activities, (b) observation, (c) medication, (d) treatment, (e) psychosocial support, and (f) rehabilitation and teaching. Indicators within each item were further divided into five categories: (a) independent, (b) minimal, (c) moderate, (d) extensive, and (e) intensive. Scores of 0 - 4 were assigned to each of the five categories, according to the amount of nursing care needed for each. Scores ranged from level A, 0 - 4, to level F, 21-24. A patient's classification level was determined by summing the scores obtained from the selected indicators under each item. Standard nursing time was calculated by using mean care hours per patient class, that is, from 1.00, 3.01, 4.60, 6.81, 10.65, to 12.8 in a sequence from class A to F. The 24-hour nursing workload of a unit was calculated in two steps. First, nursing hours needed per patient class were calculated by multiplying the number of patients in each class by the standard hours assigned per patient class. Then, the 24-hour workload for the unit was calculated by summing all the nursing hours obtained per patient class.

The Medicus system serves as the second example for the acuity-based factor evaluation/checklist style. This system contains 37 indicators, with each indicator having a different weighted score according to the amount of nursing care needed, as determined by the indicators. Scores range from type I, 0 - 24 to type V, 181 and up. The acuity level value for each class was established as Class I, 0.5; Class II, 1.0; Class III, 2.3; Class IV, 3.8; Class V, 5.5 (Medicus System Corporation [MCS], 1990). Nursing hour needed per acuity level was set hospital-wide as 4 hours. Nursing time required per patient class can be calculated by multiplying 4 hours times the acuity value per patient class. Thus, this sytem helps determine the level of care needed and the nursing time required for each patient.

Procedure-based evaluation is the second subtype of factor evaluation. This subsystem is analogous to Lewis and Carini's (1984) time or relative value unit style. In the time or relative value unit (RVU) style, the indicators used are primarily procedure-based indicators. RVU is a numerical value that represents a range of minutes needed to complete a particular task. For example, one RVU represents <6 minutes, and two RVUs represent 7 - 12 minutes to perform a procedure or task (Lewis & Carini, 1984). Standard nursing time, or RVU, for each indicator can be established by time-and-motion studies or self-reports (Dijkers & Paradise, 1986; Meyer, 1978). Some studies have shown that one nurse can provide up to 80 RVU of nursing care per shift.

In addition, total care time was determined by factoring in indirect care time, teaching and emotional support, and a delay and fatigue factor. For example, a fixed time factor, such as 10% of the total nursing hours required by the unit, was added into the unit's total required care hours (Davis & Bertram, 1991). This fixed time factor indicates time for indirect care provided and is treated as non-productive time. Therefore, each nursing unit should devise its own method of determining the time or ratio of direct, indirect, or non-nursing tasks according to its patients' characteristics, location of the unit, and philosophy of the nursing department. With accurate assessments of nursing care required by individual patients, procedure-based factor-type systems could be used as measurements for cost analysis as well as for staffing (Dijkers & Paradise, 1986).

Some systems simplify the total RVU scoring style by further categorizing RVU scores into several classes. Nursing time needed for each class is calculated in the same manner as acuity-based systems. Indirect care is treated as a fixed time, or it may not be factored in the calculation (Phillips, Castorr, Prescott, & Soeben 1992). Dijkers & Paradise (1986) suggested that this limitation may portray inaccurately the link between score and nursing care hours needed. Using this method for nurse staffing may threaten the usefulness of the instrument to determine staff needs or cost of nursing care.

The use of procedure-based factor-type systems to determine nurse staffing increases the risk of considering that the care rendered is equal to the care needed. This is an area that nurse administrators should not overlook when implementing a procedure-based system.

An example of the procedure-based factor-type system is the Grace-Reynolds Application and Study of Peto (GRASP), used in Grace Hospital (Meyer, 1978). The GRASP system contains 46 indicators, each with different score points, according to nursing tasks rendered and equal to 6.5 minutes. Total scores are determined by adding all scores obtained from nursing tasks performed, plus time needed for indirect care, teaching and emotional support, and a delay and fatigue factor. A standard time calculated for the indirect time and the time for teaching and emotional support was based on time and frequency studies. The indirect time was counted as 38 minutes, and the time for teaching and emotional support was set at 14.5 minutes. The delay and fatigue factor accounted for 12% of the total time. The scores were converted to patient care units (PCUs). Because each PCU has included indirect nursing time and fatigue factor, patient classification scores ranging from 0 - 5 can be converted to 1 PCU that represents nursing time ranging from 0 - 89 minutes (1 hour and 29 minutes). Patient classification scores ranging from 6 - 13 can be converted to 2 PCUs that represent the amount of time between 90 minutes (1 hour and 30 minutes) and 149 minutes (2 hours and 29 minutes), and so forth.

For simplicity, PCUs are rounded off to the nearest hour; for example, 0 - 89 minutes is rounded to one hour. Based on this formula for calculating patient score points to each number of PCUs, nursing workload for each unit can be determined.

In conclusion, there are many different types of patient classification systems used in hospitals. Each type has its advantages and disadvantages. The prototype approach is simple to use but is not specific because of the gross assessment of patient characteristics. A checklist acuity-based system may lack a comprehensive and systematic approach to the assessment of patient needs (Giovannetti, 1979). Time or RVU procedure-based systems are more appropriate to determine nursing costs rather than nurse staffing. A common problem in hospitals is to have a system that is limited in differentiating between the hours of care patients need and the hours of care they receive. This factor limits the application of the time or RVU procedure-based systems in projecting staffing needs.

The final outcome of classification may be the same for both prototype and factor evaluation; however, the rating procedure is different (Abdellah & Levine, 1979). In the factor evaluation method, the patient is classified by calculating a total patient classification score. The total patient classification score is the sum of all subscores assigned to the indicators for the patient, and results in a more objective status than does the prototype evaluation. Factor evaluation also encompasses a larger number of factors

in the estimation of patient needs, thereby enhancing its value as a tool for nursing care planning (Abdellah & Levine, 1979).

After weighing the advantages and disadvantages of each type of classification tool, the checklist or acuity-based factor-type system was judged to be the most efficient and effective method for patient classification if a comprehensive and systematic approach is followed in developing the indicators for assessing patients' needs. By using a nursing model as a theoretical framework in the development of a system, this instrument should reflect the complex spectrum and domains of nursing practice and patient care.

Levels of Patient Classification

The first patient classification systems that were developed categorized patients into three levels (Wolfe & Young, 1965). Due to the progressiveness of high technology and the advances experienced in health care delivery during the past few decades, patients with multiple trauma, critical illnesses, and other life-threatening situations are maintained and survive due to the interventions of the health care team. This progressiveness has led to the need for additional categories of patient classification to reflect nursing care requirements (Arenth, 1985; Roehrl, 1979; Van Der Walt, 1992). Thus, the numbers of the classification levels became an important aspect in instrument development. Current systems range from three to six levels (Des Ormeaux, 1977; Nagaprasanna, 1988;

Van Der Walt, 1992), and are usually indicated by Roman numerals I through III or VI (Des Ormeaux, 1977) in an ascending scale from minimal care requirements (I) to intensive care requirements (VI) (Hoffman, 1988).

Characteristics of Hospital Units

To classify patients accurately, indicators of care must be developed and must be both simple and efficient (De Groot, 1989b). A single list of indicators that encompasses the care activities or patient needs in all specialties will become too complex and time-consuming to complete accurately; therefore, different indicators are necessary for different types of specialty areas.

For a PCS to meet the standards of simplicity and efficiency, while reflecting nursing workload requirements to meet patient needs for each particular unit, it must include only those indicators that capture the distinctive nature and characteristics of those particular patients. Therefore, several systems may be used within a given hospital, a different one in Obstetrics, Rehabilitation, Pediatrics, Critical Care, Medical/Surgical Unit, and so on (Slyck, 1991).

The advantage of using one system throughout the entire hospital is that the criteria are equally applied to all nursing units. For example, all patients in Class II consume the same amount of nursing care resources (Slyck, 1991); therefore, hospital staff could compare one unit to another. However, the disadvantage is that the appropriateness of one system for

different specialty units that have varying acuity levels and patient needs is highly questionable and threatens the validity of the findings.

A master system would need to be developed to coordinate the classification system in a hospital, based first on one narrowly focused classification system, such as the medical-surgical system, and then expanded to include and weight all the essential indicators used for all units. Thereafter, the indicators for the specialty unit would be tailored for the specific patient population. The resulting instrument for each specialty unit would consist of the most-often used indicators adopted from the master list. Using this method, the data generated would reflect more accurately the hospital-wide patient needs and nursing requirements.

In conclusion, the complexity of types of systems varies greatly. To develop a patient classification system, four aspects must be taken into consideration: methods of indicator measurement, types of patient classification system, levels of patient classification, and characteristics of the hospital units. The acuity-based factor-type system was judged to be the most efficient and effective method for patient classification if a comprehensive and systematic approach is followed in developing the indicators for assessing patients' needs. With the increase in patient acuity level, numbers of patient classes must be extended to reflect more efficiently the required nursing workload per patient class. The characteristics of a hospital unit is an unneglectable point for the

development of a patient classification. Nurse administrators must be familiar with these four aspects and know the advantages and disadvantages of each prior to developing, modifying, or implementing a PCS in their hospitals.

Frameworks Used in the Development of Patient Classification Systems

Haas (1988) advised that the indicators used in any instrument should be based on nursing theory. Only with such a theoretical framework is it possible to define the domain of requirements for nursing care time. Through a review of the literature, two articles were found that reported use of a nursing model for the development of a PCS (Donnelly, 1981; Leatt et al., 1981). There is a tendency to identify only those symptoms and activities perceived to represent accurately nursing care requirements (Niemeier & Reed, 1985).

Giovannetti (1978) stated that a typical system includes indicators with some observable nursing services. Given the nature of the nursing function, nursing services may embrace the activities associated with feeding, bathing, ambulation, incontinent care, and activities related to medical treatment, such as preoperative preparation, observations, and special therapy. This list of nursing activities may come from empirical testing of lengthy lists of nursing activities and patient symptoms. For example, Chagnon et al. (1978) developed a procedure-based system with

the help of 100 experienced nursing staff to select 129 indicators from a long list of all nursing interventions. The indicators were grouped into nine categories: hygiene and physical comfort, feeding and hydration, elimination, breathing, supervision, patient activity, teaching, therapy, and participation in diagnostic procedures.

Hoffman and Wakefield (1986) conducted an extensive review of factor-type systems and found that almost all of them contained similar general assessment categories describing elements of nursing care and patient activities of daily living (ADLs) (Katz & Akpom, 1976). The nursing care categories usually include nutrition, intake/output, vital signs, bath/skin care, ambulation, medications, treatment, teaching, psychosocial, respiratory, and special considerations (i.e., hearing/visual loss, confusion, isolation).

Those indicators developed from Katz and Akpom's (1976) index of ADL and selected nursing activities lacked comprehensiveness and provided limited attention to some components of nursing care, such as psychosocial support and health teaching. Thus, it was suggested that classification systems should be based on the theoretical framework of a nursing model (Giovannetti, 1978; Haas, 1988).

Nursing Theory Used as a Framework

In a search of the literature, only two instances were reported where a nursing theory was used as a framework for system development. An

instrument using Roy's theory was used in one and Orem's model was used in the other.

Donnelly (1981) noted that traditional systems, which focus only on hygiene, nutrition, elimination, activity, medication administration, and vital sign monitoring, often inadequately assess patient needs. She suggested that Roy's adaptation nursing model provided a theoretical framework for the effective development of a system and demonstrated a sample of a prototype approach that contained seven components: physiological, psychological, sociological, cultural, spiritual, patient teaching, and discharge planning. The first five components of the system were derived from Roy's model and the remaining two components were added to reflect current nursing practice. Tests for reliability and validity of the instrument were recognized as important steps for validation of the instrument; however, there were no findings presented nor is the actual implementation of the instrument reported.

There have been several studies related to the development of instruments derived from Orem's model. However, only one instrument was found to be a patient classification system in which Orem's model was used as part of its theoretical framework.

Leatt et al. (1981) developed a five-level, acuity-based, factor-type instrument for assessing and classifying chronically ill patients. She used a health model to identify the main elements or events important to the

assessment of these patients. Accordingly, four main categories of elements identified from the literature were used as the conceptual framework: patients' demographic characteristics, physical status, psychosocial status, and self-care practices. Using these four factors as a basic framework, the instrument consisted of 12 major elements (i.e., physical status, mental functioning, personal and social adaptability, quality of life, activities of daily living, and therapeutic measures, and so forth). Among the 12 elements, adaptability and therapeutic measures were identified from Orem's model (Orem, 1971). Adaptability refers to patients' capability to cope with their illness, disability, or prescribed regimen. Therapeutic measures included patients' skill and judgment in carrying out prescribed treatments, management of medications, ability to contact health professional assistance, and the need for supervision in managing treatments. Others came from Katz and Akpom's (1976) index of ADL, behavioral disorders in the elderly, and other sources. The resulting initial instrument consisted of 137 variables/items.

The initial instrument was sent to a group of health care personnel that included physicians, nurses, social workers, and some consultants involved in long-term care research and to 30 patients in long-term facilities to assess the clarity of the form. The final instrument consisted of 82 items and 130 variables. By using this instrument, patients could be classified into Class I (room and board with supervision) through Class V (acute

convalescent stage). The reliability of the instrument was tested using the overall percentage agreement on patient classes as rated by health practitioners and a criterion team within the agency, resulting in 57% agreement. This outcome was somewhat disappointing. No further data related to validity testing of the instrument were reported. In addition, the instrument is complex and focuses on long-term care and, thus, it is unlikely to be used on a routine basis to classify patients in an acute-care general hospital.

Of the two systems that used nursing models as theoretical frameworks for the development of the instrument, Roy's model was used in one and Orem's model was used as a part of the framework in the other instrument. None of the instruments identified were based primarily on a framework of Orem's model. Following a review of nursing models in the literature, Orem's model was identified by the researcher to be an appropriate framework for the development of a patient classification system. Therefore, a patient classification instrument developed from Orem's self-care theory was deemed an area in need of study.

Testing of Patient Classification System

Because systems are designed primarily to determine nurse staffing, the validity and reliability of the instrument becomes the most critical factor for success in implementation to reflect patient needs accurately (Giovannetti & Mayer, 1984; Ruman & Nelson, 1987). Before 1981, few

studies reported on the reliability and validity of instruments (Leatt et al., 1981). Many systems failed because nursing service leaders had no means of verifying that the day-to-day implementation of patient classification was being done accurately (Giovannetti, 1979). During the past decade, greater attention has been directed toward testing for reliability and validity (Ambutas, 1987; Giovannetti & Mayer, 1984; Schroder et al., 1986; Williams, 1988). The following section describes methods used by hospitals in checking the reliability and validity of PCSs.

Reliability

Reliability is the prerequisite of validity. Unless patients are classified in a consistent fashion, the instrument is not useful (Williams, 1988). Many hospitals determine reliability by using interrater reliability (Jennings, Rea, Antopol, & Carty, 1989). As an additional check of instrument reliability, some hospitals use the method of correlation coefficient, a test-retest method using vignettes approved by the institution (Helberg, 1989; Verran, 1986). Haas (1988) suggested the use of multiple interrater reliability checks of the instrument. Giovannetti and Mayer (1984) recommended that the sample size used for the test be 15-20% of a unit's census, or five patients, whichever is greater. Horowitz (cited in Washington & Moss, 1988) advocated rating a minimum of 10 patients under each category, grouped by certain specific criteria in which the underlying theoretical construct was measured, such as sex or risk category.

Statistical methods used in checking the interrater reliability of the instrument in the literature included simple percentage agreement, kappa, and correlation coefficient.

Simple percentage agreement has been used widely in inpatient hospitals, because it is relatively easy to calculate, it is considered sufficient to identify areas most in need of revision, and it can provide an estimate of the extent of the problem (McKenzie, 1991). However, the opportunity for high reliability values may occur by chance alone, and there is no opportunity to correct the inflated values.

Suen and Ary (1989) identified two ways by which percentage agreement is increased by chance. First, agreement due to chance increases as the number of categories decreases. For example, chance agreement per category for a three-category system is 33%, whereas the opportunity for chance agreement per category for a five-category system is only 20%. Second, the interrater percentage agreement may be inflated by chance alone if most of the indicators are either frequently selected or infrequently selected. Accordingly, when using a simple percentage of agreement to establish reliability, a high value of results is required.

Reports from the literature suggested that agreement of 90% or more is acceptable (Giovannetti & Mayer, 1984; Haas, 1988; Huckabay & Skonieczny, 1981). Giovannetti and Mayer indicated that, 80% to 90% may need to be discussed to clear up some misunderstanding, and agreement

below 80% indicates the need for further education or a review of the rules for classifying, or both. Shelley (1984) indicated that reliability above 90% requires excessive time and strict standards for administration. The resulting scale or test might then become too time-consuming and complex to administer in a research setting. She suggested that an 80% rater agreement should be reached for instrument reliability, and when tested for research, reliability as low as 60% is adequate. In addition, a study published by Churness, Kleffel, Onodera, and Jacobson (1988) considered a minimum acceptable standard of 65% agreement between nurses and auditor ratings.

Kappa is a reliability value that reflects the degree of agreement between two raters after eliminating agreement accounted for by chance alone (Soeken & Prescott, 1986; Woolson, 1987). Thus, the use of kappa provides a more accurate reflection of interrater reliability. Although the computation of kappa is more complex than percent agreement, appropriately programmed computers simplifies the application of this statistic (Jennings et al., 1989).

A kappa value of zero means that observed agreement equals chance agreement. Kappa's upper limit of 1.00 would indicate perfect agreement, whereas a negative kappa would indicate observed agreement is less than that expected by chance. Landis and Koch (1977) identified the strength of agreement according to six divided ranges of the kappa statistics:

1. Strength of agreement is poor when kappa <0.00.

2. Strength of agreement is slight when kappa ranges from 0.00 - 0.20.

3. Strength of agreement is fair when kappa ranges from 0.21 - 0.40.

4. Strength of agreement is moderate when kappa ranges from 0.41 - 0.60.

5. Strength of agreement is substantial when kappa ranges from 0.61 - 0.80.

6. Strength of agreement is perfect when kappa ranges from 0.81 - 1.00.

Haas (1988) recommended that at least an 0.9 correlation on types of classification should be achieved by the raters before the training session is considered successful. An example of testing reliability of PCSs using the statistical method of correlation coefficient was conducted on a pediatric hematology oncology unit (Lovett, Wagner, & Mcmillan, 1991). To test the correlation value of the instrument, the agreement between the level of patient care was evaluated by two independent raters on a sample of 150 patient observations. The statistical method used was the Pearson product moment correlation coefficient and resulted in a high correlation (r = .97; p <.001). However, no correlation based on patients' total scores was reported in this study.

An acceptable reliability takes several months to achieve and is only possible after an appropriate in-service education program for the raters (Giovannetti, 1979). For example, after 2 months of advanced training and practice, the percentage agreement on indicators used for patients who had end-stage renal disease increased from 63% to 95% (Kessler et al., 1990).

When an acceptable reliability is reached, periodic checks should be made to ensure that reliability continues. Alward (1983) suggested that a weekly or at least a monthly auditing process should be implemented to monitor interrater reliability of the system. De Groot (1989a) pointed out that reliability should be monitored on a quarterly basis and this action should be included in part of the Nursing Quality Assurance Program description. In Van Der Walt's (1992) study, reliability was checked every 3 months.

De Groot (1989a) suggested the use of a patient constant approach in the testing of instrument reliability. In this approach, nurses classify patients based upon some hypothetical patient situation or set of patient characteristics that are documented in the nurses' notes, Kardex, or other written documentation. Poulson (1987) cautioned against using written profiles because of the tendency to read them hurriedly and miss obvious information, or read into the profile information that has not been stated specifically. Thus, it was suggested that written documentation should not be substituted for checking interrater agreement on actual patients.

Validity

Validity refers to how accurately the system reflects actual nursing care requirements. Although validity is considerably more difficult to evaluate than reliability (Jennings et al., 1989), it is important to evaluate periodically at least the validity of the average minutes of care per activity or per class (De Groot, 1989a). In 1987, Ambutas identified three important types of validity: content, criterion-related, and construct validity. Among these three methods, predictive validity was applied primarily to establish criterion-related validity. Not many instruments were found that had been tested for construct validity. It was the aim of this research to demonstrate the construct validity of the instrument by using the method of contrasted groups construct validity. The following will describe three types of validity testing used in the study: content validity, contrasted groups validity, and predictive validity.

Content Validity

Demonstrating content validity is the first task in establishing the accuracy of a data collection instrument (Thomas, 1990). According to Giovannetti (1979), "content validity has no empirical basis and relies generally on judgment" (p. 7). Therefore, content experts selected for verifying the content of the instrument should be conversant with the domain treated in the measuring tool.

Usually, two or more content experts are employed for one study. The researcher should develop appropriate criteria to guide the selection of content experts. One or more poor content evaluators can greatly compromise the process of content validation (Waltz et al., 1991). Further, two qualified content experts are better than three poor content evaluators. It is difficult to establish the content validity for PCSs, because no specific guidelines were found in the literature to assist in testing the validity of these systems (Alward, 1983).

To check content validity of the instrument, the most frequently used approach is the use of content specialists to assess the quality and representativeness of the items within the test for measuring the content domain (Waltz et al., 1991). The index of content validity (CVI) can be used to quantify the extent of agreement between the experts. For example, the relevancy of each item to the objective(s) is determined using a 4-point rating scale: not relevant, somewhat relevant, quite relevant, and very relevant. The CVI is the proportion of items that give a rating of quite/very relevant by both experts involved.

Contrasted Groups Construct Validity

Most variables that require development of a measure with properties of reliability and validity are constructs (Shelley, 1984). The contrasted groups technique is the most common approach to test construct validity of instruments (McLaughlin & Marascuilo, 1990). At present, few

systems using the contrasted groups method to test their construct validity were found in the literature.

To test the contrasted groups construct validity, the investigator identifies at least two groups of individuals who are known to have significant differences in the characteristic being measured by the instrument (Waltz et al., 1991). The instrument is then administered to the groups to obtain data for examination. If the instrument is sensitive to individual differences in the trait being measured, the mean performance of these groups should differ significantly using a statistical procedure such as t-test or ANOVA.

Predictive Validity

Predictive validity is a subtype of criterion-based validity and is different from concurrent validity, the other subtype of criterion-based validity. Predictive validity refers to evidence of a relationship between the new data collection instrument and a criterion in the future. Concurrent validity seeks to establish that the performance of one measure is correlated with the performance on another measure or criterion. If the classification performance of one PCS has concurrent validity, a high correlation on the performance of another valid PCS would be expected. However, concurrent validity of a PCS is difficult to establish due to the paucity of valid patient classification systems (Alward, 1983).

Predictive validity can be tested by determining whether the instrument classifies patients in a way that agrees with the amount of nursing care that must be provided. Predictive validity, then, is used for validation of the patient classification system, not just of the instrument (Williams, 1988). De Groot (1989a) identified two ways of establishing predictive validity: establishment of the average amount of care associated with a given patient class and establishment of standard care time for individual nursing activities, treatments, or procedures.

Alward (1983) indicated two basic approaches to determine time required to provide care for each class of patients: a time-and-motion observation study and self-recording of activities. Thompson and Diers (1991) suggested the use of consensus on time allocation to set time standards. Methods such as time-and-motion studies are expensive and intrusive. When they are not feasible, the nurses' perception of nursing time needed by patients or the consensus of nurses on time standards per patient class can be pursued as alternative mechanisms for decision-making.

To evaluate the subjective estimations of nursing time needed per patient class in comparison with actual time used, Fries and Cooney (1984) examined comparisons made between the nurse's subjective estimation and objective estimation of time spent caring for 426 patients in five nursing homes. A Pearson correlation of 0.20 with a significant value at the 0.001 level was found when comparing nursing time estimated by nurses and

actual nursing time spent. A Pearson correlation of 0.11 with a significant value at the 0.05 level was found when comparing estimated time spent by aides and actual time spent by aides. A Pearson correlation of 0.20 with a significant value at the 0.001 level was found again when comparing total nursing time (average of aide and nurse) estimated and actual.

To ameliorate the weakness of subjective measures on adequacy of nursing time, Williams (1988) suggested a gross but realistic approach in testing predictive validity of a PCS, which is to ask staff members about the number of patients in each class they could care for adequately on a particular shift. The average of the estimates is calculated and then tested by assigning that number of patients to a staff member. The adequacy of the assignment is evaluated at the end of the shift.

In summary, a PCS needs to be valid to have credibility (Ambutas, 1987). There are many different tests used to validate the instrument. When developing a PCS, the developer must evaluate carefully the reliability and validity of the instrument using appropriate and feasible methods. The high degree of validity may be rewarded by a correspondingly accurate costing of the use of valuable nursing resources (Hay & Nelson, 1988).

Training Program

The most important task in developing and implementing a system in a hospital is adequate training of nurses to classify patients (Chagnon et al.,

1978). Errors in patient classification may be due to insufficient knowledge of patients' needs and conditions or a lack of knowledge about the system (Niemeier & Reed, 1985). Classifying patients without proper training in the use of the instrument may result in misclassifications and can be discouraging to the nurses. Inadequate training can result in rejection of the system, one of the primary reasons for failure experienced in the implementation and applications of PCSs (Giovannetti, 1979).

A well developed training session should consist of at least four stages:

1. general discussion in groups, explaining objectives, structure, content of the classification, and operation of the system;

2. practical experience of classifying in small groups of four or five persons, based upon patient observation and chart documentation;

3. pooling of results in the large group and discussion of difficulties and disagreements; and

4. individualized training for persons who experience special difficulties (Chagnon et al., 1978).

In addition, a user's manual or guide that contains a detailed description of the system, abbreviations, and definitions should be developed and studied by those who are using the system.

In the training of staff, Poulson (1987) suggested the use of a set of written patient profiles based on actual patients who have been

hospitalized. The profiles should be tested for ambiguity before use by having two experts independently classify the patient described. After modification, the profiles should be retested by other experts and continue until agreement on all profiles reaches a satisfactory level. A goal of 90% on patient type and 80% on use of the indicators for the set of profiles should be established for each nurse before data collection begins.

The training sessions should be started 2 to 3 weeks prior to the start of data collection (Poulson, 1987). A 2-hour group training session is useful so individuals may raise questions that benefit the entire group. Some hospitals use computer programs to orient staff to the classification system. Experience in using the computer program indicates that it is an effective and efficient method of training, only requiring about 30 minutes (Gender, 1989). Other hospitals use audiotape instruction and have found this method to be effective with minimal training time (Schroder et al., 1986).

Influence of Computers

Computers were first used for patient classification in the 1960s to assist in calculating staffing needs for a nursing unit (Wolfe & Young, 1965). Today, computer use has expanded to include monitoring information related to productivity, budget planning, trend analysis, costing, charging, and linking of the system to a variety of patient data, such as

quality criteria, length of stay, nursing diagnoses, and medical care data (Giovannetti & Johnson, 1990).

The interface of the computerized system with a computerized staffing program enables the staff to determine patient acuity levels and display the number of personnel needed on the next shift for each nursing unit. Other uses for the interfaced systems would be to generate several management reports that aid in auditing and monitoring the system, such as patient days per classification, unit staffing schedules, nursing utilization (productivity), and monthly unit acuity summaries (Arenth, 1985; Gender, 1989). In some hospitals, the patient classification system program is linked with the billing system program through a mainframe computer. The daily nursing care charges may be determined according to the patients' acuity-levels and simultaneously transferred to the billing system program (Stepura & Miller, 1989).

Another use of the computer permits retention of data generated from PCSs, thereby making it relatively easy to track for reliability and validity on an ongoing basis. The traditional approaches of monitoring reliability and validity are time-consuming and costly, and they frequently lead to the abandonment of reasonable classification instruments (Giovannetti & Johnson, 1990). Additionally, basing nurses' training on computer-generated results of PCSs saves time and money (Gender, 1989).

Computers speed-up the development and implementation of PCSs. In Nagaprasanna's (1988) survey of 213 hospitals, 54% reported they used a computerized patient classification system. Having the ability to classify patients in a computer program saves nursing and clerical time, facilitates audits, and improves the ability to generate management reports (Gender, 1989). Computer use for developing and implementing systems, as well as for determining their validity and reliability of the system on a periodic basis, will become an indispensable step in the future.

Comparisons Among Systems and Timing of Classification

Roehrl (1979) conducted a 9-week study in which three 4-level classification systems were compared in a medical unit and a surgical unit. Two systems used the check-list acuity-based classification methods, one the procedure-based. Patients were classified once per day from 2:00 p.m. to 7:00 p.m. by registered nurses: the head nurse, and the day and evening team leaders. A valid sample of 779 showed the percentage agreement by patient types between the two tools to be as high as 64%. The percentage agreement by patient types/classes among the three tools was 34%. The author suggested that some modification of the point ranges in the third tool would greatly enhance the percentage agreement with the other tools. The study did not address interrater reliability for the three tools.

In 1982, Jackson and Resnick compared two instruments. The first was a 4-level prototype with 28 acuity-based indicators used in Montefiore Hospital Medical Center (MHMC), in which the interrater reliability ranged from 75% to 100% agreement. The second instrument, the Therapeutic Inventory Scoring System (TISS), also a 4-level factor PCS, was composed of 75 procedure-based indicators. The investigators reported that TISS had proved useful and reliable worldwide; however, no reliability of the instrument was described in the original report by Cullen, Civetta, Briggs, and Ferrara (1974). Each indicator on the tool was assigned 1 - 4 points, ranging from the least to the most amount of time required. Results showed 68.2%, or 90 out of 132 classification differences, by patient type. Therefore, the researchers questioned whether any system that had not been tailored for an institution's circumstances could be useful.

Ambutas (1988) compared the Rush-Medicus 4-level system (RMPCS) with TISS in an intensive care unit at George Washington University. RMPCS is a 4-level factor system with 32 acuity-based indicators. Each indicator is assigned a certain weight for determining nursing workload. Before data collection began, 40 medical intensive care staff nurses (an all RN staff) attended a 1-hour in-service class describing proper use of the RMPCS and TISS tools. Interrater reliability between the investigator and staff nurses was found to be 91.6% using RMPCS, and 81.6% using TISS. The correlation coefficient between RMPCS and TISS

scores was r = .67, (p, <.0001; N = 307). Feedback solicited from data collectors indicated that neither tool measured all activities necessary to provide patient care. TISS is task-oriented and does not consider the psychosocial needs of patients and family. Likewise, RMPCS does not allow for many therapeutic needs performed in the intensive care unit. One recommendation from the study was to continue investigating an ideal system for classifying patients in intensive care units.

The frequency of classifying patients per patient day differs from system to system. The literature reports instances where patients are classified three times a day (Niemeier & Reed, 1985; Strom, 1993), twice a day (Gender, 1989; Van Der Walt, 1992), or once a day (Dowman, Spence & Sankaran, 1990; Helmer & Halbert, 1987).

Kinley and Cronenwett (1987) reported a study based on a convenience sample of 621 cases that used the GRASP-based system to classify patients once for each of three shifts. A paired t-test and the correlation coefficient were used to determine whether there were significant differences among patient care unit (PCU) (1 PCU = 1 hour) totals by shift. Results showed that there were no significant differences among the three shifts in the nursing workload (t = 0.05 - 1.10, p >.05; r = .95 - .99, p <.0001). They suggested that in a hospital with an average patient census of 350, a savings of 6,300 hours of RN time per year could be realized by classifying patients once instead of three times a day. In

consideration of a balance between cost-effectiveness and accuracy of the system, the investigators recommended that one daily measurement of patient care requirements would be the most appropriate method for the hospital. No interrater reliability was reported in the study.

Batty et al. (1990) examined the differences in unit workloads across three shifts using Rush-Medicus PCS (RMPCS) in a general surgical unit. The 39 nurse raters maintained a 90% reliability during the entire 28-day study period. Data were collected from patients at 10:00 a.m., 6:00 p.m., and 2:00 a.m. each day of the study. The power spectral analysis, analogous to ANOVA, on a total of 2,843 patient observations showed no remarkable differences in the mean unit workload across the three shifts. The evenly distributed workload on all shifts might be attributed to situations such as same day admission surgical, short stay chemotherapy, and emergency admissions that added to the acuity on the evening shift of the unit. In fact, on an average, the evening and night shifts tended to have higher acuities than did the day shift. This finding suggests that unit managers should evenly distribute the nursing staff over the three shifts.

In the same study, Batty et al. (1990) suggested that patients be classified on all three shifts, or for the classification to be updated at least during the evening or night shifts. When the patients were classified for each of the three shifts during the 28 days of the study period, the units were found to be understaffed for 13 days and it was then possible to

allocate staff accordingly. When the patients were classified once per day (at 10:00 a.m.) for allocation of nursing staff, only 7 of the 28 days were found to be understaffed. The latter finding is judged to be less sensitive and, thus, less accurate. The insensitive results on a daily measurement of patient classification and the lack of a clear pattern in the daily workload make it difficult to allocate nursing staff adequately.

In an intensive care unit, instruments were used once per each of the three shifts to classify patients for nurse staffing purposes (Niemeier & Reed, 1985). Classifying patients from all three shifts is believed to yield the most accurate data because it captures changes in patient status (Kinley & Cronenwett, 1987). The fluctuation of patients' condition, especially in intensive care units, supports the use of more frequent classifications to make more accurate staffing decisions for patient care. Additionally, with the use of a computerized classification program, the additional time required to classify patients is negligible (Gender, 1989).

To classify patients twice a day, instruments usually are administered once on the day shift and once on the evening shift. For example, the day staff have until 11:00 a.m. to rate patients so that staffing can be predicted for the evening shift. The evening staff completes ratings by 7:00 p.m. to use for determining night and day shift staffing (Gender, 1989).

When classifying patients once a day, the results serve as a basis to predict staffing for the next three shifts. The time of day selected to

classify patients varies from instrument to instrument. For example, some classifications may be performed at 10:00 a.m. (Cottey et al., 1992; MSC, 1990) or at 12:00 a.m. (midnight) (Arenth, 1985). Others may occur about noon (Helmer & Halbert, 1987) or be classified sometime during the evening shift (Kinley & Cronenwett, 1987).

In Nagaprasanna's (1988) survey on those who used 4-level classification systems, 60% of the respondents answered that they classified patients daily, whereas 40% classified patients once on each of the three shifts. Regardless of which method used, it may take several months' experience and comparisons among several PCSs before a decision can be made about the appropriate frequency for classifying patients.

Time needed for classifying a patient also differs from system to system. Niemeier and Reed (1985) reported that independent raters spend less than 10 minutes classifying 12 patients using a TISS factor evaluation tool in an intensive care unit. Poulson (1987) reported that raters using the Rush-Medicus tool took approximately an hour to assess three to five patients on a general unit.

Time required for completing a patient's classification may be related to the familiarity of the nurse raters with the instrument and the ease of classification. The extra time needed for classifying patients by nurses unfamiliar with the instrument may be spent in checking the classification methodology and definitions of indicators. Even nurses who are familiar

with the system need periodic clarification of indicator descriptions to assure accuracy. For ease of use, the instrument should be unambiguous without overlapping categories (Alward, 1983), conceptually clear and concise, and as objective as possible to avoid subjective interpretation (De Groot, 1989b).

The traditional method of using patient census as the primary factor in the determination of nurse staffing is no longer suitable. Instead, patient classification systems provide a useful method to measure nurse staffing requirements. In examining the trend in the development and implementation of systems, it is clear that there are many different types used in hospitals, yet many have been abandoned, changed, or improved to provide more useful information (Davis & Bertram, 1991).

Summary

This chapter reviewed the development of patient classification systems used in hospitals. Types of systems were analyzed, and the acuity-based factor-type was identified as the best method for patient classification. The need for a well planned training program that is designed to teach classification skills for staff nurses is deemed essential. The influence of computers was identified. None of the PCSs reviewed were recognized as entirely satisfactory for measuring nursing needs and predicting nurse staffing needs. The lack of conceptual or theoretical frameworks in the development of PCSs was discussed, and the few that

have been used were identified. Orem's self-care model was judged by the researcher as an appropriate framework to use for the development of a patient classification instrument in the clinical setting.

CHAPTER III

METHODOLOGY

The purpose of this study was to develop a six-level acuity-based factor type patient classification system (PCS) for medical-surgical patients between 18 and 65 years of age, using a theoretical framework derived from Orem's (1991) self-care model, and to pilot test the instrument's reliability and validity in an acute-care teaching hospital in the southeastern region of the United States. The future goal is to validate and implement this newly developed acuity-based system in the medical-surgical units of an acute care teaching hospital in Taiwan. This newly developed system, the Orem Patient Classification System (OPCS), consists of (a) 18 items that reflect patient needs, (b) indicators that reflect nursing care needed, (c) patient classification form, (d) scoring system, and (e) Table for the Estimation of Nursing Workload Based on Patient Class.

This chapter describes the research design and hypotheses, setting and population, research steps of the study, development and revisions of the instrument, and training of the raters. The formulation of the PCS and methods used in testing the validity and reliability of the instrument are also presented.

Research Design

This research was designed to answer the question "What is the support for reliability and validity of a patient classification instrument that is derived from Orem's self-care model for use in an acute-care teaching hospital?" Consequently, this methodological study included a combination of descriptive, factorial, and correlational research design. The nursing workload for each patient class was established based on a prediction equation obtained from data gathered in the pilot study.

Research Hypotheses

Three research hypotheses were examined in the study. The significance level for all three hypotheses was set at alpha = .05. Hypotheses 1 tested interrater reliability of the instrument, Hypothesis 2 tested for contrasted groups construct validity, and Hypothesis 3 tested for predictive validity. All the hypotheses were written in both research and null forms.

Hypothesis 1

Research: There will be a significant interrater agreement on items of the OPCS among the nurse raters (interrater reliability).

Null: There will be no significant interrater agreement on items of the OPCS among the nurse raters.

Hypothesis 2

Research: Patient total classification scores for the OPCS will demonstrate a high to low sequence among patients in medical-surgical intensive care units, patients not ready for discharge from medical-surgical units, and patients ready for discharge from medical-surgical units (contrasted groups construct validity).

Null: There will be no difference in patient total classification scores among patients in medical-surgical intensive care units, patients not ready for discharge from the medical-surgical units, and patients ready for discharge from medical-surgical units.

Hypothesis 3

Research: Using the OPCS, there will be a positive correlation between patient total classification scores and the nursing time required by patients (predictive validity).

Null: Using the OPCS, there will be no correlation between patient total classification scores and the nursing time required by patients.

Setting and Population

The setting for this study included five medical-surgical units, one medical intensive care unit, and one surgical intensive care unit in an acute-care teaching hospital in the southeastern region of the United States. Prior to the collection of data from observations of hospitalized patients and chart reviews, the study proposal was approved by the Nursing

Research Committee of the hospital and by the Institutional Review Board for Human Use Approval of The University of Alabama at Birmingham. Both groups determined that the research protocol presented no risk to the subjects who participated.

The study population included patients from five medical-surgical units, one medical intensive care unit, and one surgical intensive care unit. The subjects were drawn randomly from this population, were within the age range of 18 - 65 years, and had been hospitalized on those units for more than 8 hours prior to data collection. Additionally, intensive care unit patients who were not ready for transfer to general units at the time the instrument were used for classification. The study sample was selected to test the instrument for interrater reliability, contrasted groups construct validity, and predictive validity.

Development of the Instrument

Information for the development and testing of the instrument included reviews of literature related to nursing theories and patient classification systems, consultations with experts in instrumentation, and visits to several teaching hospitals to assess the use of existing patient classification systems in measuring nursing workload. Through these series of investigations and consultation activities, a patient classification instrument derived from Orem's self-care model for measuring patient

acuity was developed. The nine steps used to develop and test the instrument were as follows:

1. Reviewed the literature and gathered additional information from experts in PCSs, and from several teaching hospitals that had implemented a patient classification system.

2. Developed the first draft of the instrument based on Orem's self-care requisites and included the development of items, categories, indicators, scoring system, and directions and guidelines for the use of the instrument.

3. Assessed content validity using content experts. The appraisals from the experts were measured by calculating a Content Validity Index (CVI).

4. Revised the instrument using feedback from the content experts and CVI results.

5. Assessed content validity for the revised instrument using the same content experts and a second CVI value was calculated.

6. Conducted training sessions for registered nurse raters on the use of the instrument.

7. Field-tested the instrument for interrater reliability by percentage agreement estimation, using a small sample of six hospitalized patients.

8. Repeated the testing of the instrument for interrater reliability; tested for contrasted groups construct validity and predictive validity assessment, using a sample of 30 hospitalized patients.

9. Established a six-level classification system for the medical-surgical Orem's Patient Classification System.

In summary, steps 1-5 related to the development of the instrument, step 6 related to the training of the raters, steps 7-8 related to field-testing the instrument, and step 9 related to the establishment of the classification system. These nine steps are summarized in Figure 2.

Orem's Self-Care Requisites

A valid patient classification system should have the power to identify patient's self-care deficits and determine nurse staffing in a cost-effective way. To assure that the patient classification system was appropriate for measuring nurse workload, the theoretical framework based on Orem's self-care model served as the basis for constructing the items and indicators for the instrument. Self-care requisites derived from Orem's (1991) theory of self-care served as the building blocks in the construction of instrument items. The concepts of self-care agency, self-care deficits, and nursing agency served as the framework for organizing the indicators under each item. Both the items and indicators were developed within the framework of universal, developmental, and health-deviation self-care requisites. Each

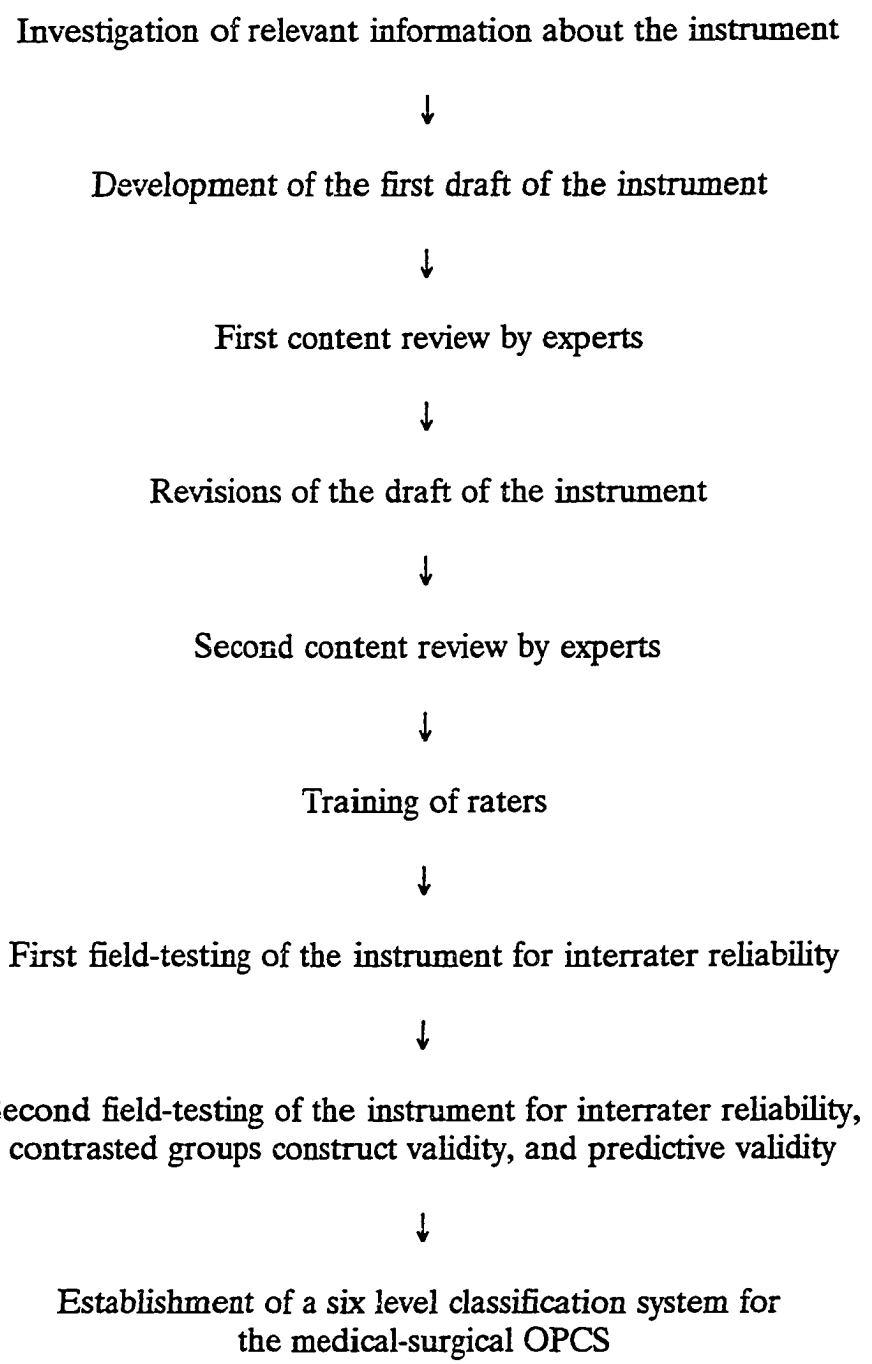


Figure 2: Steps in development and testing the OPCS

type of requisite represents a category of deliberate actions to be taken by or for individuals because of their needs as human beings. The three self-care requisites of universal, developmental, and health-deviation are described in the following paragraphs.

Universal Self-Care Requisites

Universal self-care requisites are associated with life processes and maintenance of the integrity of human structure and function. Requisites in this category include: (a) sufficient intake of air, water, and food; (b) care for elimination processes and excrements; (c) maintenance of a balance between solitude and social interaction; (d) maintenance of a balance between rest and activity; (e) prevention of hazards to life; and (f) promotion of normalcy.

Developmental Self-Care Requisites

Developmental self-care requisites were initially subsumed under the universal self-care requisites. They were separated to emphasize their importance and diversity (Orem, 1991). Developmental self-care requisites are associated with human developmental processes and conditions or events that could adversely affect development or could prevent the occurrence of adverse effects. Thus, requisites associated with human developmental processes encompass various stages of the life cycle, such as intrauterine stages of life, process of birth or neonatal stage of life, infancy, childhood, adolescence, and adulthood. Conditions and events that

influence the developmental processes may include status of education, social adaptation, status of health, relatives and friends, possessions and occupational security, living conditions, and environment of residence.

Health-Deviation Self-Care Requisites

Health-deviation self-care requisites are associated with defects or pathological changes that require nursing regulation or treatment. Examples of requisites in this category include genetic and constitutional defects, human structural and functional deviations and their effects, and treatment measures prescribed or performed by physicians. Orem (1991) also identified six categories of health-deviation self-care requisites, such as seeking and securing appropriate medical assistance, and being aware of and attending to the effects and results of pathologic conditions. Health-deviation self-care requisites demand the use of valid and reliable measures to make practical judgments about self-care deficits in order to meet the patients' therapeutic self-care demands and restore the patients' self-care agency.

Development of the Items

A review of Orem's model shows that the self-care requisites and basic conditioning factors delineated in the theory of self-care can serve as the building blocks for the construction of items for a patient classification instrument. Self-care requisites include items such as the maintenance of sufficient air, water, and food. Basic conditioning factors contain items

influence the developmental processes may include status of education, social adaptation, status of health, relatives and friends, possessions and occupational security, living conditions, and environment of residence.

Health-Deviation Self-Care Requisites

Health-deviation self-care requisites are associated with defects or pathological changes that require nursing regulation or treatment. Examples of requisites in this category include genetic and constitutional defects, human structural and functional deviations and their effects, and treatment measures prescribed or performed by physicians. Orem (1991) also identified six categories of health-deviation self-care requisites, such as seeking and securing appropriate medical assistance, and being aware of and attending to the effects and results of pathologic conditions. Health-deviation self-care requisites demand the use of valid and reliable measures to make practical judgments about self-care deficits in order to meet the patients' therapeutic self-care demands and restore the patients' self-care agency.

Development of the Items

A review of Orem's model shows that the self-care requisites and basic conditioning factors delineated in the theory of self-care can serve as the building blocks for the construction of items for a patient classification instrument. Self-care requisites include items such as the maintenance of sufficient air, water, and food. Basic conditioning factors contain items

whereas other terms, such as air, water, and elimination processes, were explicit and could be used easily in item construction. Those terms that needed to be defined are as follows:

1. Rest: emotional stabilization of the individual.

2. Prevention of hazards to life: integrity of the sensory system of the individual.

3. Activity: strengths or restrictions of the neuromuscular and skeletal system of the individual.

4. Promotion of normalcy: compliance to the therapeutic regime or acceptance of the individual as in need of care.

5. Processes of development: age-relevant stages of life processes of the individual that include the neonatal stage, childhood, adolescence, and adulthood.

To eliminate the possibility of ambiguity or overlap among categories (Alward, 1983), concepts thought to be similar or closely related among the three self-care requisites in Orem's model were integrated into one item. Concepts thought to have more than one focus were broken down into subconcepts and used to generate different items. Concepts thought not applicable to medical-surgical patients between 18 and 65 years of age were excluded.

Examples of integration of concepts included: terminal illness and impending death (Orem's developmental self-care requisites) and poor

health and disability (Orem's developmental self-care requisites). Terminal illness and impending death were believed to be related to the concept of pain, which was an indicator under universal self-care requisite items. Poor health and disability were believed to be among the health-deviation self-care requisites.

Examples for concepts that were broken into subconcepts and used to generate different items included nursing care for elimination processes and maintenance of the balance between activity and rest from Orem's universal self-care requisites. The concept of care for elimination processes was considered to have two foci, care for both urinary and bowel systems. Thus, subconcepts and items developed from this concept were considered to be bladder elimination (physical need) and bowel elimination (physical need). The concept of maintenance of the balance between activity and rest had two foci, physical mobility and ability to maintain psychological comfort. Thus, subconcepts and items developed from this concept were physical motion (physical need) and comfort (psychological need).

Concepts that were deleted because they were not applicable were developmental stages (Orem's developmental self-care requisites) and seeking and securing appropriate medical assistance (associated with human pathology in Orem's health-deviation self-care requisites). Developmental life stages were omitted because the classification instrument was to be used only with the young and the middle-age adult patients. Seeking and

securing appropriate medical assistance was omitted because the instrument was to be used with hospitalized individuals.

A first draft of the instrument was developed and consisted of 19 items (see Appendix A). Among these, 15 were considered mandatory items (applicable to all patients) and 4 were considered optional (applicable only in some patient situations). The 19 items are reviewed below.

Items 1-9 were developed to represent the area of universal self-care requisites: air intake, fluid intake, food intake, bladder elimination, bowel elimination, prevention of hazard, activity, rest, and promotion of normalcy. Items of air intake, fluid intake, food intake, and promotion of normalcy were drawn directly from Orem's framework of self-care requisites. Items of bladder elimination and bowel elimination were derived from the framework of care associated with eliminative processes and excrements. Items of activity and rest were drawn from the framework of maintenance of a balance between activity and rest.

Items 10-12 were categorized to represent the areas of developmental self-care requisites: health education, social support, and financial security. The item of health education was drawn directly from the framework of developmental self-care requisites. The item of social support resulted from the integration of concepts in both universal and developmental self-care requisites. The balance between solitude and social interaction came from the universal self-care requisites, and the item relating to relatives,

friends, and associates came from the developmental self-care requisites. Financial security is associated with possessions or occupational security and was drawn from developmental self-care requisites.

Items 13-15 were categorized as health-deviation self-care requisites: lab data collection, medical regimen, and surgical regimen (dressing changes). Health-deviation self-care requisites arise because of an individual's disabilities in self-care or from medically prescribed measures to prevent pathology or to compensate for disability (Orem, 1991). The individual's disabilities in self-care requisites were identified as the deficits in universal and developmental self-care requisites and were disseminated under items of these two categories. Therefore, only three health-deviation self-care requisite items were developed: laboratory data collection, medical regimen, and surgical regimen.

Four additional items, categorized as optional, identified patients' extra disabilities in the area of health-deviation self-care requisites. They were previously identified as deficits in universal and developmental self-care requisites and were disseminated under items of these two categories. The four optional items, weighted more heavily because of their influence on and demand for intensive nursing manpower, could only be applied to patients with the conditions as described by the indicators. Nurses would rate these four items only if the patient being classified required tracheal or endotracheal suctioning at least once every four hours, had multiple

intravenous (IV) lines, required frequent vital signs or intake and output check, or had orders for special examinations or treatment procedures.

In summary, the first draft of the instrument consisted of 19 items, 15 mandatory and 4 optional items. They included:

1. air intake,
2. fluid intake,
3. food intake,
4. bladder elimination,
5. bowel elimination,
6. hazard prevention,
7. activity,
8. rest,
9. normalcy promotion,
10. health knowledge,
11. social support,
12. financial security,
13. lab data collection,
14. medical regimen,
15. surgical regimen,
16. frequent use of tracheal or endotracheal suctioning,
17. multiple IV lines,
18. frequent vital signs, intake and output checks, and
19. special examination or treatment procedures.

Development of the Indicators

Indicators for the instrument were derived from three concepts of Orem's model: self-care agency, self-care deficit, and nursing agency. In other words, the individual's state is expressed in terms of self-care agency, self-care deficits, or nursing agency needed.

Self-care agency, the main concept from the theory of self-care, is the capability of regulating one's own functioning and development, such as the ability to breathe unassisted, have regular bowel movements, and have

adequate range of motion. Patient self-care agency was primarily expressed in three types of indicators, patient conditions (e.g., regulate bowel movements), patient capability (e.g., ambulatory), and lack of identified self-care deficit (e.g., no respiratory difficulty). The main concept from the theory of self-care deficit has to do with an individual whose self-care agency is insufficient to meet therapeutic demands, such as the need for oxygen or regular intravenous infusion. Patient's self-care deficits were expressed in patient conditions (e.g., anxiousness) or lack of identified self-care agency (e.g., urinary incontinence). Nursing agency, the main concept from the theory of nursing system, denotes the ability of nurses to meet an individual's therapeutic demands, such as tracheal suctioning and tube feeding. Nursing agency was expressed as nursing tasks performed (e.g., total feeding), and treatment regimens (e.g., use of colostomy). When indicators for each item are framed by the main concepts from Orem's three nursing theories, the range of nursing workload can be determined systematically.

A group of indicators was developed initially by the researcher for each item based on a review of 10 randomly selected patient charts in the acute-care teaching hospital. Some indicators were explicit; others were more abstract and needed to be defined. Therefore, a list of definitions for indicators was also developed to provide the necessary information for the nurse raters. To increase clarity, the indicators and definitions were

reviewed by an expert in instrumentation, an expert in nursing theory, and experienced volunteer nurses on separate versions. Unclear, overlapping, or redundant indicators and definitions were modified. A few indicators were added, and some wording was changed following feedback from the experts.

Based on the capacity of self-care agency, the degree of self-care deficits, and the extent of required nursing agency, the indicators under each item were further categorized according to minimum, moderate, and intensive nursing care needs. Feedback from the two experts and the three volunteer nurses aided the process of categorizing the indicators. A score of 1 was assigned to the category "needing minimum nursing care," 2 to "moderate need for nursing care," and 3 to "need for intensive nursing care."

A weighing system was established by the addition of several new optional items to measure patients' extra disabilities in the area of health-deviation self-care requisites or to identify additional nursing agency required by the patients. These items required more nursing manpower than others. For example, the item of "maintenance of sufficient intake of air" should include indicators such as "no respiratory difficulty," which would fall in the category of minimum nursing care needed. Indicators such as "use of chest tube" and"tracheal or endotracheal suction x 1-2 shift" would fall in the category of moderate nursing care needed. Indicators

such as "tracheal or endotracheal suction 3-4/shift" and "use of ventilator" would fall in the category of intensive nursing care needed. The optional item created for maintenance of sufficient intake of air was "frequent use of tracheal or endotracheal suction." Indicators under this new optional item included "tracheal or endotracheal suction x 5-6/shift" in the category of minimum nursing care, "tracheal or endotracheal suction x 7-8/shift" in the category of moderate nursing care, and "tracheal or endotracheal suction x 9/shift or more often" in the category of intensive nursing care. By using this method, patients' self-care deficits can be quantified and weighted, and nurse staffing needs can be calculated and determined.

Development of the Patient Classification Form

The final activity in developing the draft of the OPCS instrument was the formulation of directions and guidelines, and the construction of a patient classification form. Directions provided general descriptions of the instrument, processes for patient classification, and guidelines for nurse raters who classified patients. The Patient Classification Form (PCF) was used by nurse raters to record data for the classified patient. Data included on the PCF were: (a) the maximum number of patients similar to the subject being classified that the rater could care for per shift, (b) patient demographic data, and (c) individual item scores and total cumulative score. Following the completion of the first draft of the OPCS instrument, it was subjected to content evaluation.

Testing Instrument for Validity and Reliability

The purpose of this research was to develop a medical-surgical, acuity based patient classification system for patients who were between 18 and 65 years of age, using a theoretical framework derived from Orem's self-care nursing model and to pilot test its reliability and validity in an acute care hospital. A series of criteria relative to quality and utility guided the process of instrument development. These criteria related to issues of reliability and validity are discussed in subsequent sections.

Content Validity

Validity refers to the degree to which a measure is capable of achieving the purposes for which it was developed (Waltz & Bausell, 1986). The validity tests of an instrument can be broadly categorized into three types: construct validity, criterion-related validity, and content validity (Ambutas, 1987). Construct validity is determined by using one or more of four methods: contrasted groups, hypothesis testing, multitrait-multimethod, and factor analysis. Criterion-related validity is determined by using one or more of two methods: concurrent validity and predictive validity (Thomas, 1990).

Content validity refers to the adequacy with which a specified domain of content is sampled (Ebener, 1985). It focuses on the representative nature of a group of items related to the specified content domain (Waltz et al., 1991). As Nunnally (1978) noted,

> if it is agreed by most potential users of the test, or at least by persons in positions of responsibility, that the plan was sound and well carried out, the test has a high degree of content validity. (p. 92)

This study selected several methods that were considered to be sufficient in testing the validity of the instrument being developed, including content validity, contrasted groups construct validity, and predictive validity. To proceed with the testing, a panel of experts representative of the field of study were deemed appropriate to provide an evaluation of the suggested items for this instrument and to suggest changes to improve its focus. The instrument was tested for content validity twice before it was deemed appropriate to continue the interrater reliability, contrasted groups construct validity, and predictive validity testing.

Two approaches were used to measure content validity of the instrument: a Content Validity Procedure and an Item Content Assessment Sheet (see Appendix B). The Content Validity Procedure provided the content experts with procedures used for measuring content validity of the instrument. The Item Content Assessment Sheet contained items to be evaluated according to the following three questions:

1. Is the item congruent with the concept of self-care requisites presented by Orem (1991)?

2. Is the item relevant to the objective of the instrument for measuring the nursing staff workloads for patient care?

3. Does the item specifically measure a self-care requisite, the domain of interest under each mode?

A 4-point rating scale was used for these questions. Items were rated for congruency, relevance, and specificity on a 4-point scale, ranging from 1 = not congruent, 2 = somewhat congruent, 3 = quite congruent, and 4 = highly congruent on the item content assessment sheet. This rating scheme was also applied to questions 2 and 3 evaluating the relevance and specificity for each item. When the mean item rating across the two content experts was ≤ 2, the item was modified or deleted. Several blank lines were provided at the end of the Item Content Assessment Sheet for recommendations from the content experts.

For the first testing of content validity, two content experts evaluated the instrument. One was a recognized expert in Orem's model, the other was an expert in nursing service administration. They examined and rated each item independently.

To quantify the extent of expert agreement on the instrument's validity, an index of content validity (CVI) was calculated (Waltz et al., 1991). The CVI was the proportion of questions rated as quite/very congruent, quite/very relevant, or quite/very specific for each item by both raters involved. A CVI of .80 is considered an acceptable level of content validity (Waltz et al., 1991). Revisions of the instrument and the content validity assessments were continued until the CVI for the items was $\geq .80$.

The percentage agreement (P_o) between the two experts on each item was also tested. The P_o was defined as the proportion of given items rated as one or two plus items rated as three or four by both experts (Waltz et al., 1991). Table 1 presents the results of the first evaluation of the instrument on content validity using the CVI and the P_o of the CVI value between the two experts on each item. The CVI value for the congruence of items with Orem's theory was .79, which was considered unacceptable; the CVI on relevance to staffing was .95; and the CVI on specificity to self-care requisites was .16, suggesting that only 16% of the items were considered by both content experts to be quite/very specific. The specific results were below an acceptable point. The percentage agreement among expert raters was only .21, indicating the two content experts disagreed on the specificity of these new items.

The instrument was revised based on the feedback obtained from the first testing for content validity. The six revisions are described as follows:

1. The financial security item was deleted.

2. Seven items were modified to highlight the domain of nursing interests. These included items 6 and 7, physical motion from activity and comfort from rest, and items 11-15, social relations from social support, information monitoring from laboratory data, delivery of medicine from medical regimen, skin care from surgical regimen, and frequent tracheal suction from frequent use of tracheal or endotracheal suction.

Table 1

First Content Validity Index (CVI) and Percentage Agreement (P_o) of Items Rated by Two Content Experts

	Expert 1				
Expert 2	Rated 1 or 2	Rated 3 or 4	Total items	CVI	P_o
	Congruence to Orem's theory				
Rated 1 or 2	0	0	0		
Rated 3 or 4	4	15	19		
Total items	4	15	19		
CVI				.79	
P_o					.79
	Relevancy to nursing staffing				
Rated 1 or 2	0	1	1		
Rated 3 or 4	0	18	18		
Total items	0	19	19		
CVI				.95	
P_o					.95
	Specific to self-care requisites				
Rated 1 or 2	1	0	1		
Rated 3 or 4	15	3	18		
Total items	16	3	19		
CVI				.16	
P_o					.21

3. The normalcy item was redefined as health knowledge about the disease or injury of the individual.

4. The prevention of hazard item was changed to reflect both the integrity of sensory organs (physical need) and orientation (psychological need).

5. The indicators for each item were examined. Indicators for newly developed items (integrity of sensory organs and orientation) were developed. Items number 3, 10, and 14, food intake, health knowledge, and medical regimen, were clarified by rewording or modifying some of the indicators. Additional indicators and definitions were added or modified to clarify the domain of interests.

6. Finally, the Patient Classification Form was refined and amended to update the changes made on the numbers and topics of the items.

A transformation table (Table 2) was developed to explain and clarify how each item was derived from Orem's framework. Table 2 shows the transformation processes that indicate how Orem's concept of self-care requisites/deficits converted into the related items. For example, the concept of maintenance of a sufficient intake of air, a universal self-care requisite, was transformed into the ability to take in air through the respiratory system. Therefore, the term developed for this item was air intake.

After the revision, the instrument consisted of 18 items, 14 mandatory and 4 optional items. They included:

1. air intake,
2. fluid intake,
3. food intake,
4. bladder elimination,
5. bowel elimination,
6. physical motion,
7. integrity of sensory organs,

Table 2

Transformation Table From Orem's Framework to Items of PCS

Requisites/deficits of self-care	Focus of patient/nursing agency	Item of patient classification
Universal self-care requisites common to all human beings:		
(1) Maintenance of a sufficient intake of air	(1) Ability to take in air through the respiratory system	(1) Air intake (physical need)
(2) Maintenance of a sufficient intake of water	(2) Ability to maintain appropriate amount of fluid	(2) Fluid intake (physical need)
(3) Maintenance of a sufficient intake of food (nutrient)	(3) Ability to take in nutrients	(3) Food intake (physical need)
(4) Provision of care associated with elimination processes and excrements	(4) Ability to take care of bladder elimination	(4) Bladder elimination (physical need)
	(5) Ability to take care of bowel elimination	(5) Bowel elimination (physical need)
(5) Maintenance of a balance between activity and rest	(6) Ability to move	(6) Physical motion (physical need)
	(7) Ability to maintain psychological comfort	(7) Comfort (psychological need)

Table 2 (continued)

Requisites/deficits of self-care	Focus of patient/nursing agency	Item of patient classification
(6) Maintenance of a balance between solitude and social interaction	(8) Ability of maintaining appropriate social relations	(8) Social relations (social need)
(7) Prevention of hazards to human life, human functioning and human well-being	(9) Integrity of sensory organs	(9) Sensory organs (physical need)
	(10) Integrity of cerebral function	(10) Orientation (psychological need)
(8) Promotion of normalcy	(11) Recognizing health state of self	(11) Health knowledge (social need)
Health-deviation requisites arise from the measures used in diagnosis or treatment		
(9) Health-deviation self-care requisites arise from general measures used in diagnosis	(12) Assistance with general measures used for diagnosis	(12) Information monitoring (need for medical action)
(10) Health deviation self-care requisites arise from general measures used in treatment	(13) Assistance with general measures used in treatment regimen	(13) Delivery of medicine (need for medical action)
		(14) Skin care (need for medical action)

Table 2 (continued)

Requisites/deficits of self-care	Focus of patient/nursing agency	Item of patient classification
Optional self-care deficits originating from health deviation self-care requisites		
(11) Health deviation self-care requisites from maintenance of a sufficient intake of air	(14) Ability to take in air through special system	(15) Frequent tracheal suction
(12) Health deviation self-care requisites from maintenance of a sufficient intake of water	(15) Assistance with intake of fluid by extra route	(16) Multiple IV lines (need for medical action)
Optional self-care deficits originate from health-deviation self-care requisites arising from measures used in diagnosis or treatment		
(13) Health deviation self-care requisites arise from special measures used in diagnosis	(16) Assistance with special frequency of monitoring measures	(17) Frequent VS or I&O (need for medical action)
(14) Health-deviation self-care requisites arise from special measures used in diagnosis or treatment	(17) Assistance with special procedure of diagnosis or treatment	(18) Special examination or treatment procedure (need for medical action)

8. orientation,
9. comfort,
10. social relations,
11. health knowledge,
12. information monitoring,
13. delivery of medicine,
14. skin care,
15. frequent tracheal suction,
16. multiple IV lines,
17. VS or I&O, and,
18. special examination or treatment procedures.

A cover sheet that provided a brief introduction was designed for the instrument. The instrument consisted of: directions for using the Medical-Surgical Patient Classification Instrument, listing of patient classification indicators of nursing care needed, definitions of patient classification indicators, and the Patient Classification Form (see Appendix C). The first 14 items of the revised instrument were mandatory; the last 4 were optional. Potential scores for the revised instrument ranged from 14 to 54. The instrument was considered ready for the second round of content validation.

The revised instrument was tested for content validity the second time using the Content Validity Procedure and Item Content Assessment Sheet (see Appendix D). Results of the re-evaluation are presented in Table 3.

In Table 3, all CVI values were above the set criterion of .80 (Waltz et al., 1991) and were acceptable on all three questions. Based upon the second content evaluation, minor changes were made.

Table 3

Second Content Validity Index (CVI) and Percentage Agreement (P_o) of Items Rated by Two Content Experts

				Expert 1	
Expert 2	Rated 1 or 2	Rated 3 or 4	Total items	CVI	P_o
		Congruence to Orem's theory			
Rated 1 or 2	0	0	0		
Rated 3 or 4	1	17	18		
Total items	1	17	18		
CVI				.94	
P_o					.94
		Relevancy to nurse staffing			
Rated 1 or 2	0	2	2		
Rated 3 or 4	0	16	16		
Total items	0	16	18		
CVI				.89	
P_o					.89
		Specific to self-care requisites			
Rated 1 or 2	0	1	1		
Rated 3 or 4	1	16	17		
Total items	1	17	18		
CVI				.89	
P_o					.89

Interrater Reliability

Reliability refers to the consistency with which a measurement rule assigns scores to a group of examinees (Waltz & Bausell, 1986). The reliability of an instrument can be assessed in three ways: stability, internal consistency, and equivalence (Polit & Hungler, 1991). Stability refers to the consistency of an instrument on repeated administration. An example of

stability assessment is test-retest reliability. Internal consistency refers to the extent of consistency of the instrument's subparts in measuring the same attribute. An example of internal consistency assessment is split-half reliability. Equivalence refers to two different assessments: one indicates the consistency of a measure using different observers measuring the same individual at the same time, the other, the consistency of a measure using a different instrument/form to test the same individual at the same time. An example of equivalence is interrater reliability for the former and parallel forms of the test for the latter.

The method chosen in testing an instrument's reliability depends, to a certain extent, on the nature of the instrument and the aspect of the reliability concept that is of greatest interest (Polit & Hungler, 1991). Because of the nature of PCSs, a measure of interrater reliability is the most important and widely used method for reliability testing (Giovannetti, 1979). Thus, the instrument developed for this study was tested for interrater reliability.

Interrater reliability refers to the agreement among two or more raters in assigning scores to the objects or responses being judged. In this study, three nurse raters were used to check the interrater reliability. Methods applied for the test included the percentage agreement and a one-tail z test on kappa statistic. Percentage agreement is relatively easy to calculate and is considered sufficient to identify areas most in need of

revision (McKenzie, 1991). During the training sessions and pre-testing of the instrument, tests of interrater reliability among the nurse raters were calculated using percentage agreement among raters on items of the instrument (O'Neil, 1972; Polit & Hungler, 1991; Waltz et al., 1991). Kappa is a reliability value that reflects the degree of agreement between two raters after the agreement accounted for by chance alone has been eliminated (Soeken & Prescott, 1986; Woolson, 1987). The interrater reliability on items of the instrument for the pilot study was tested by the kappa statistic in combination with a one-tailed z test. The level of statistical significance for the one-tailed z test on kappa value was set at alpha = .05 (Fleiss, 1981; Waltz et al., 1991; Woolson, 1987).

From a review of the literature, it was clear that there should be multiple interrater reliability checks of an instrument (Haas, 1988) and that patient classifications should be done by registered nurses (Roehrl, 1979). In this study, three raters were used to check the instrument: two volunteer registered nurses and the nurse researcher. The three nurse raters trained together using a program developed by the researcher. The raters used the revised instrument that had been tested for content validity to classify patients.

The two volunteer registered nurses were employed by the hospital where the testing occurred. Both were graduates of diploma programs and

had more than 10 years nursing experience in acute care settings. All three raters participated in 10 hours of interrater training.

The training program included orientation to and practice in using the instrument. Training sessions were as follows:

1. In the first two sessions, nurse raters practiced their classification skills independently, with 15 patient vignettes. These vignettes were used by the hospital to orient nurses to the hospital's patient classification system.

2. In the third session, nurse raters practiced their classification skills by classifying three patient charts selected from one hospital unit.

3. In the fourth session, nurse raters practiced classification, using six randomly drawn patient charts from the Medical Record Office.

The effect of the training program was evaluated by the test of interrater reliability. The method used for checking interrater reliability was to compare the percentage of agreement among the three raters on each item. The percentage agreement on the items was calculated using the formula of O'Neil (1972):

$$\frac{\text{Number of agreements}}{\text{Total number of agreements + disagreements.}}$$

The weakness of using patient charts in the data collection is that the chance of getting a problematic or deficient data set is rather high (Polit & Hungler, 1991). Therefore, the minimal criteria for interrater agreement

was set at 75% for chart rating, and at 80% for clinical field testings (Poulson, 1987; Shelley, 1984). The percentage agreement among raters was 64% at the beginning and 80.2% by the end of the training session. The training program was considered effective enough to test the instrument in the field.

The first field-testing of interrater reliability of the instrument included a convenience sample of six hospitalized patients, age 18 to 65 years. These subjects were hospitalized in three of the five participating medical-surgical units and verbally consented to participate in the study. The nurse raters obtained data from shift nurses' reports, reviewed patients' charts, observed the patients, and asked nurses questions about the patients' conditions. If the selected patient was conscious, the three nurse raters explained the purposes and procedure of the study to them. Occasionally, nurse raters asked patients questions regarding their condition or feelings.

The raters recorded patient data and independently scored the items on the patient classification form (PCF) using the patient classification instrument. After data collection from each patient, the nurse raters estimated the nursing time required and independently classified each patient. To estimate nursing time required by each patient, the raters evaluated the maximum number of similar patients a nurse could care for per shift and entered that number on the PCF. The assessments took into account the desired quality of nursing care that each patient should receive.

The determination of nursing time required for the patient was computed as 8 hours per shift divided by the number of patients estimated. A total of six patients were classified during the field test.

The percentage of agreement among items scored by the raters was calculated to test interrater reliability of the instrument (O'Neil, 1972). A total of 324 pairs of scores (3 pairs x 18 items x 6 patients) was obtained and compared. Raters agreed on 278 of the 324 pairs of ratings, resulting in 89.1% agreement (278 divided by 324), which exceeded the criteria of .80 set by Shelley (1984). Thus, the instrument was considered ready for the second field test using a large sample of 30 hospitalized patients. The estimate of nursing hours was intended for use only as a reference.

The second field-testing of the instrument occurred in five medical-surgical units and two intensive care units and included testing for interrater reliability, contrasted groups construct validity, and predictive validity. For the second field-testing of the instrument, a stratified random sample of 30 patients was selected from five medical-surgical units, a medical intensive care unit, and a surgical intensive care unit. The same criteria were used for selection of the subjects as used for the first field test. Stratified random sampling procedures were conducted to obtain patients in all care categories (Giovanetti & Mayer, 1984). The 30 patients selected were as follows:

1. Group I consisted of 10 patients from the intensive care units, 5 from a medical intensive unit, and 5 from a surgical intensive care unit. The patients selected were not ready for transfer to general units on the classification day.

2. Group H was made up 10 patients identified by unit nurses as not ready to be discharged within 48 hours. Two patients were selected from each of the five medical-surgical units.

3. Group D included 10 patients identified by unit nurses as ready for discharge within 48 hours from the same five medical-surgical units as group H. Two patients were selected from each unit.

Data collection procedures were similar to those in the previous testing of the instrument. Nurse raters collected patient demographic data, estimated patient conditions, judged the required nursing time, and independently classified the patient. These actions were repeated until all of the 30 patients were classified using the PCF and the patient classification instrument.

A descriptive research design was applied to test the interrater reliability of the instrument. The expected agreement (*Pe*) generated for each item and for the cumulative totals served as the control groups for kappa (*ka*) statistics. The kappa value obtained for each item and for the total measure were compared with those of expected agreement. The formulas for these statistics were as follows:

$$P_o=\frac{1}{NK(K-1)}[(\sum_{i=1}^{N}\sum_{j=1}^{R}n_{ij}^2)-NK]$$

$$P_j=\frac{1}{NK}\sum_{i=1}^{N}n_{ij}$$

$$P_e=\sum_{j=1}^{R}P_j^2$$

$$K_a=\frac{P_o-P_e}{1-P_e}$$

$$SE_k=\sqrt{\frac{2}{NK(K-1)}\frac{\sum_{j=1}^{R}P_j^2-(2K-3)(\sum_{j=1}^{R}P_j^2)^2+2(K-2)\sum_{j=1}^{R}P_j^3}{(1-\sum_{j=1}^{R}P_j^2)^2}}$$

$$Z=\frac{K_a}{SE_k}$$

where K_a = kappa, the chance-corrected proportion agreement among items,

P_o = observed proportion of agreement among items,

P_e = expected proportion of agreement among items due to chance alone,

P_j = the proportion of all classifications that fall into group j,

N = total number of patients classified,

K = number of raters per patients,

R = number of categories possible,

n_{ij} = number of nurse raters who rate patient i into category j.

Kappa tests the proportion of agreement among raters after chance agreement on items is removed (Fleiss, 1981). In other words, Cohen's kappa (*Ka*) excludes chance agreement and permits the calculation of a true percentage agreement among raters (Soeken & Prescott, 1986; Woolson, 1987). From the above formula, it can be seen that the statistic kappa controls for interrater agreement by chance. A kappa value of zero means that observed agreement equals to chance agreement. Kappa's upper limit of +1.00 would indicate perfect agreement. A negative kappa value indicates observed agreement less than that expected by chance.

Because 30 patients were selected for this test, it was assumed that the level of agreement (such as observed agreement, expected agreement, or kappa) generated on each item obtained from the 30 patients had a standard normal distribution. Thus, the difference between value of kappa (*Ka*) and value of expected agreement (*Pe*) on each item as well as on total items/measure of the instrument was examined using a one-tailed *z* test at alpha = .05 level. The independent variable for this test was the type of raters.

The research hypothesis for instrument testing was: There will be a significant interrater agreement on items of the OPCS among the three nurse raters. The one-tailed *z* test was to examine whether the agreement in the kappa value was statistically significant.

Contrasted Groups Construct Validity

Construct validity is an important type of validity to examine the accuracy of an instrument (Burns & Grove, 1987). The contrasted groups approach is a commonly used method of testing the construct validity of an instrument (Waltz et al., 1991). Therefore, the contrasted groups construct validity was deemed as an appropriate method for this study in testing the construct validity of the classification system.

The construct used in this study was the patient's self-care deficit. The 30 patients used to test the instrument for interrater reliability were further divided into the three previously mentioned 10-patient groups for contrasting: (a) Group I, 10 patients from the two intensive care units identified as having the greatest self-care deficits and who were not ready to be transferred to general units on the classification day; (b) Group H, 10 patients identified as having moderate self-care deficits and not ready for discharge from the medical-surgical units within 48 hours; and (c) Group D, 10 patients identified as having the least self-care deficits and who were ready for discharge from the medical-surgical units within 48 hours.

Because there were three nurse raters and three patient groups involved in this test, a 3 x 3 factorial research design using a two-way ANOVA test was used. The test of two-way ANOVA was done by using the SPSS statistical software package. Statistical significance was set at $p \leq .05$ value. Analysis by two-way ANOVA checked effects for both of the

patient groups (the first independent variable or the first main effect) and the nurse raters (the second independent variable or the second main effect) on patient total classification scores (the dependent variable). Patient total classification scores denoted scores that were summed from each item of the instrument for a patient.

Additionally, this two-way ANOVA test also checked whether there was an interaction effect of the nurse raters and the patient groups on patient total classification scores. The alpha value was set at .05 level. If the result of the two-way ANOVA showed there was an interaction effect between the patient groups and nurse raters, the effect of patient groups on patient total classification scores would be separately analyzed for each of the three raters using the statistics of one-way ANOVA and Scheffé's tests. If there is no interaction between the two variables ($p > .05$), results on both first and second main effects from two-way ANOVA test can be interpreted independently.

If both main effects can be independently interpreted, a $p \leq .05$ value obtained from the variable of patient group (the first main effect) statistically signifies the differences on mean patient total classification scores among the three patient groups, which could, in turn, support the contrasted groups validity of the instrument. A $p > .05$ obtained from the variable of nurse rater (the second main effect) indicates that there were no significant differences on mean total patient classification scores among

these three nurse raters. This homogeneity of the nurse raters could further support the interrater reliability of the instrument.

If results from the two-way ANOVA showed that (a) there was no interaction effect between variables of patient groups and nurse raters ($p > .05$), (b) the variable of patient groups did affect patient total classification scores ($p \leq .05$), and (c) the variable of nurse raters had no effect on total patient classification scores ($p > .05$), then it would be considered that the patient group was the only factor that had an impact on the patient total classification scores and that the difference among nurse raters had no influence on patient total classification scores. Subsequently, other statistics may be applied to test the contrasted groups construct validity of the instrument.

If the patient group was considered as the only factor that affected patient total classification scores, data for the three nurse raters within each patient group would be combined. Accordingly, a cluster of 30 (10 patients x 3 raters) patient total classification scores was included in each patient group. These reconstructed data were tested by one-way ANOVA and the Scheffé test using the SPSS statistical software package. It was this one-way ANOVA test that provided a chance for the application of the special follow-up test called Scheffé test. The independent variable was patient group, and the dependent variable was patient total classification scores.

The alpha value for both one-way ANOVA and Scheffé tests was set at .05 level.

If the results from the one-way ANOVA showed a significant difference among the three patient groups, a Scheffé's method of multiple comparisons would be applied to determine which patient group contained patient total classification scores different from those of other groups. In addition, the mean patient total classification scores among these three patient groups was also calculated for comparison. If the instrument had contrasted groups construct validity, patient total classification scores of the OPCS would demonstrate a high-to-low sequence among patients in medical-surgical intensive care units, patients not ready for discharge from medical-surgical units within 48 hours, and patients ready for discharge from medical-surgical units within 48 hours (research hypothesis 2).

Predictive Validity

Predictive validity is a subtype of criterion-related validity. Criterion-related validity is assessed by correlation of the instrument with a criterion measure. In other words, criterion-related validity involves the relationship between one measure and another measure of the same phenomenon (Woods & Catanzaro, 1988). Usually the criterion is a second measure, which examines the same concept under study.

Both predictive validity and concurrent validity are subtypes of criterion-related validity (Thomas, 1990). Concurrent validity refers to

evidence of a relationship between the new data collection instrument and an existing criterion of the same concept. An example of concurrent validity would be a correlation value of a newly developed PCS with an existing PCS. Predictive validity is similar to concurrent validity, but deals with a criterion in the future. In other words, if an instrument had predictive validity, it has the ability to predict a future outcome from the current phenomena measured. An example of predictive validity would be a correlation value between classification scores obtained from an instrument and the nursing hours needed for the patients being classified.

Because it is difficult to accept the major premise that another classification instrument is valid in other settings, testing the validity of a PCS based on its concurrent validity is open to question (Giovannetti, 1979). In addition, comparison of PCSs across hospitals involves the assumption that different classification systems are comparable in terms of their estimates, although PCSs differ greatly in a number of ways from one another (Ambutas, 1988; Jackson & Resnick, 1982; Phillips, Castorr, Prescott, & Soeben, 1992). The assumption of comparability may be unwarranted (Phillips et al., 1992). The test of concurrent validity might be feasible only if the hospital chooses to use its existing PCS as a criterion to check the validity of the PCS that has been somewhat modified from its original PCS. Therefore, this study used selected nursing time needed for

each patient being classified as a criterion to measure the predictive validity of the instrument.

To test the predictive validity of the instrument, a correlational research design was applied. Patient total classification scores were correlated with nursing care time required to determine the predictive validity of the instrument. Patient total classification scores served as the independent variable (predictor variable). The required nursing time was a dependent variable (outcome variable). Nursing time required was estimated based on the judgement of the nurse raters. If the instrument had predictive validity, patient total classification rating scores using the OPCS would be positively and significantly correlated with the estimated nursing time (research hypothesis 3).

A scatterplot was used to show how the patient total classification scores were distributed in relationship to nursing time required. Thereafter, the appropriate correlation coefficient was used to check the relationship on these two variables. A positive correlation coefficient above .70 is considered acceptable (Fox, 1982), and above .85 is considered high (Catanzaro, 1988).

Development of Scoring System

Patient total classification scores for the instrument ranged from 14 to 54. To create a six-level PCS, five cutoff points were used to divide the scores into six segments, according to the following procedures:

1. Determine the nursing time range for each patient class. The overall nursing time ranges for the instrument were determined according to nursing-time data collected from the 30 patients estimated by the three nurse raters in the field pilot study. The nursing time range for each of the six patient classes was determined by dividing the overall nursing time range by six.

If the distribution of nursing care time required for the 30 patients was found to range from 1 - 7 hours, the overall nursing-time range for the instrument would be 6 hours (7 hours - 1 hour). The subrange of required nursing care time for each of the six patient class would be 1 hour (6 hours [overall range of nursing care time] ÷ 6 [number of classification levels]). Consequently, five cutoff points that divided the overall nursing care time into six subranges or segments would be 2, 3, 4 ,5, and 6 hours: Class I patients needed from 1 - 1.9 nursing hours; Class II, from 2 - 2.9 hours; and so forth.

2. Determine the range of patient total classification scores for each patient class. According to previously used data on the correlation coefficient between patient total classification scores and required nursing care time, a regression line was drawn and the prediction equation was calculated using SPSS.

The corresponding points on patient total scores were calculated from each of the cutoff points of nursing care item using the formula of

prediction equation. The five corresponding points categorized the span of patient total classification scores into six segments and determined the score range for each classification interval of the instrument.

3. Nursing hours per patient class were determined by measuring the mean nursing care time for each of the six pre-determined nursing time ranges using data from the 30 hospitalized patients. In this way, a rudimentary six-level patient classification system was formulated.

Summary

The purpose of this study was to develop a six-level acuity-based medical-surgical PCS derived from Orem's (1991) self-care nursing model and to pilot test it for reliability and validity. The reliability assessed for this instrument was interrater reliability. Validity was assessed for this instrument using content validity, contrasted groups construct validity, and predictive validity. The instrument was first evaluated by content experts and revised until the CVI reached an acceptable level. Nurse raters were trained to achieve at least 75% agreement on items of the instrument. The first field-testing of the instrument was conducted using a convenience sample of six hospitalized patients and resulted in an acceptable percentage agreement among nurse raters.

The second field-testing of the instrument was conducted using a sample of 30 hospitalized patients randomly selected from five medical-surgical units and two intensive care units. Data collected for the second

field-testing were used to evaluate the instrument for interrater reliability, contrasted groups construct validity, and predictive validity. The nursing workload per patient classification was determined. The instrument development was considered complete when the five cutoff points that classify patients' nursing care requirement levels were determined.

CHAPTER IV

FINDINGS

This chapter presents the findings from the development and testing of a patient classification instrument in the context of an acute-care health setting. The focus of the chapter relates to the testing procedure and data analysis for demonstrating the reliability and validity of the instrument. The purpose of this study was to develop a six-level acuity-based factor-type medical-surgical PCS (OPCS) using a conceptual framework derived from Orem's (1991) nursing model and to pilot-test its reliability and validity in an acute-care teaching hospital in the southeastern region of the United States. The future goal is to further validate and implement this newly developed OPCS in medical-surgical units at an acute-care teaching hospital in Taiwan.

In addition to the development of the OPCS and demonstration of content validation of the instrument, this study included a combination of descriptive, factorial, and correlational research design. The research question for the study was, "What is the support for reliability and validity of a patient classification instrument that is derived from Orem's model for use in an acute-care clinical setting?" In responding to the research

question, three hypotheses were proposed to field-test the instrument for interrater reliability, contrasted groups construct validity, and predictive validity. Data collected from the clinical setting were used to test these research hypotheses.

Description of the Sample

The sample included 30 hospitalized patients randomly selected from an acute-care teaching hospital in the southeastern region of the United States. These 30 patients were selected from three different groups, with each group consisting of 10 patients. Group I included 5 patients selected from the medical intensive care unit and 5 patients from the surgical intensive care unit. Group H included 10 patients who were not to be discharged within 48 hours from five medical-surgical units, with 2 patients from each unit. Group D included 10 patients who were to be discharged in 48 hours from the same five medical-surgical units as group H, with 2 patients from each unit.

All 30 patients selected were alert and verbally consented to participate in this study. Data were collected from 10:30 a.m. to 03:30 p.m. on 2 different days, with a 5-day interval in between. Fourteen patients were rated on the 1st day: 4 in Group I, 4 in Group H, and 6 in group D. Sixteen patients were rated on the 2nd day: 6 in Group I, 6 in Group H, and 4 in Group D.

Table 4 presents the characteristics of the 30 patients. Patients' ages ranged from 18 to 63 years, with a mean age of 41.5 years, and a standard deviation of 14.9. Fourteen patients were 40 years of age or younger, and 16 were over the age of 40. Nineteen of the patients were male, and 11 were female; 23 were white, 7 were black; 19 were from the medical area, and 11 were from the surgical area. The compiled data supported the heterogeneity of the group and, thus, it was deemed as an appropriate sample to test the instrument.

Table 4

Characteristics of the Patients

Variable	*No.*	*%*	*M*	*SD*
Age (18-63)			41.5	14.9
≤ 40	14	46.7		
> 40	16	53.3		
Gender				
Male	19	63.3		
Female	11	36.7		
Race				
White	23	76.7		
Black	7	23.3		
Clinical Area				
Medical	19	63.3		
Surgical	11	36.7		

$N = 30$

Presentation of the Findings

Findings from the data analysis are presented in relation to the three tests: interrater reliability, contrasted groups construct validity, and predictive validity. Each test consisted of the related hypothesis and results from statistical analysis.

Interrater Reliability

The interrater reliability of the instrument was tested based on comparisons of scores obtained from items of the instrument among the three nurse raters. The hypothesis and results from the interrater reliability test are described below.

Research Hypothesis 1

The research hypothesis stated: "There will be significant interrater agreement on items of OPCS among nurse raters." The corresponding null hypothesis stated: "There will be no significant interrater agreement on items of OPCS among nurse raters." If the instrument is reliable, the results will support the research hypothesis.

Results From Statistical Analysis

Scores on items measured by the three nurse raters were first tested by the expected agreement and the kappa statistic. Subsequently, a one-tailed *z* test was applied to test the difference between the expected agreeement and the kappa statistic generated. The level of statistical

significance for the one-tailed z test on kappa value was set at alpha = .05 level (Fleiss, 1981; Waltz et al., 1991; Woolson, 1987).

Expected agreement and kappa. When all the items rated are consistent among the three raters, the agreement value would be 1.00 (100% agreement). Table 5 presents results of the observed agreement (P_o), the expected agreement due to chance alone (P_e), kappa (K_a), standard error of kappa (SE_k), and the one-tailed z value for each item as well as the total items on the instrument. Table 5 shows that the observed agreement (P_o) among each of the 18 items ranged from .74 to .97 and the overall P_o was .87. Using Shelley's (1984) criteria for testing the interrater reliability in research, this result (P_o = .87) exceeded the acceptable standard of 80%.

The expected agreement (P_e) among each of the 18 items varied from .34 to .96. The overall P_e of the instrument revealed .34. Apparently, some of the items possessed a ratio of chance agreement much more than can be expected. Those items demonstrated a skewed distribution, which made this instrument less sensitive to discriminate the differences among patients.

Kappa values among the 18 items ranged from -.01 to .95. The overall kappa statistic for the instrument was .81, exhibiting an acceptable strength of agreement among raters. Most of the kappa statistics were capable of discerning the degree of differences between P_o and P_e; the closer the P_o value to P_e value, the smaller the kappa value. For example, the kappa value in item 7 (sensory organ) was .19, reflecting little effect of

Table 5

One-Tailed Z Test of Kappa Value on Items of OPCS

Items	P_o	P_e	K_a	SE_k	Z
1. Air intake	.96	.45	.92	.11	8.68**
2. Fluid intake	.84	.37	.75	.09	8.83**
3. Food intake	.83	.44	.70	.11	6.55**
4. Bladder elimination	.97	.57	.95	.14	6.97**
5. Bowel elimination	.89	.59	.73	.16	4.60**
6. Physical motion	.93	.35	.90	.08	11.37**
7. Sensory organ	.82	.78	.19	.27	0.73
8. Orientation	.97	.83	.82	.32	2.67*
9. Comfort	.76	.34	.62	.08	8.05**
10. Social relations	.84	.60	.61	.16	3.71**
11. Health knowledge	.78	.50	.55	.19	4.67**
12. Information monitoring	.93	.36	.90	.08	11.17**
13. Delivery of medications	.78	.36	.65	.08	7.71**
14. Skin care	.86	.66	.61	.19	3.17**
15. Frequent tracheal suction	.96	.96	-.01	.69	-0.02
16. Multiple IV line	.86	.50	.71	.13	5.45**
17. VS or I&O	.96	.67	.86	.20	4.32**
18. Special exam/ treatment	.74	.39	.58	.09	6.30**
N = 30					
Total item	.87	.34	.81	.02	40.97**

Note. P_o = observed agreement
Pe = expected agreement
SEk = standard error of kappa
Ka = kappa
**$p<.001$
*$p<.01$

a high P_o (.82) from a high P_e (.78). The kappa value in item 6 (physical motion) was .90, reflecting a greater effect of a high P_o (.93) from a low P_e (.35). The kappa in item 15 (frequent tracheal suction) was -.01, indicating that the observed agreement was somewhat less than the agreement by chance alone, although both P_o and P_e for that item showed comparably high results (.96). To determine if the reliability of the instrument tested by kappa statistic reaches a level of statistical significance, a z test of kappa statistic is further applied.

Z test of kappa statistic. To check the significance level of the kappa statistic for each item, kappa obtained from each item was compared with its own P_e, using the statistic of one-tailed z test. The z values obtained for each item were compared to the critical value (1.65) at alpha = .05 level. Consequently, 16 out of the 18 items were found to have value >1.65, exhibiting significant agreements in their kappa statistic. The significant agreement among the nurse raters did not occur in two items. The z value was .73 for item 7 (sensory organ) and -.02 for item 15 (frequent tracheal suction). The overall kappa value was highly significant when compared with the overall P_e (.34) using a one-tailed z test ($z = 40.97, p<.001$). This result was consistent with the overall observed agreement (87%).

Because the overall kappa statistic was significant, the null hypothesis was rejected and the research hypothesis was supported. However, findings from the z tests on items of the instrument suggested that contents from

item 7 (sensory organ) and item 15 (frequent tracheal suction) required revisions if the instrument is to be equally reliable for all items.

Contrasted Groups Construct Validity

To test the contrasted groups construct validity of the instrument, three patient groups were identified and tested for difference. The hypothesis and results from the contrasted groups construct validity are described below.

Research Hypothesis 2

Research hypothesis number two stated, "Patient total classification scores of the OPCS will demonstrate a high-to-low sequence among patients in medical-surgical intensive care units, patients not ready for discharge from medical-surgical units within 48 hours, and patients ready for discharge from medical-surgical units within 48 hours." The corresponding null hypothesis stated: "There will be no difference in patient total classification scores among patients in medical-surgical intensive care units within 48 hours, patients not ready for discharge from the medical-surgical units, and patients ready for discharge from medical-surgical units within 48 hours."

Results From Statistical Analysis

The descriptive statistics on patient total classification scores among the patient groups and nurse raters are shown as in Table 6. With three nurse raters classifying 10 patients for groups I, H, and D, each patient

Table 6

Descriptive Statistics on Patient Total Classification Scores Among Patient Groups and Nurse Raters

Category	*N*	*M*	*SD*	Mini-mum	Maxi-mum
Patient Group I					
Nurse 1	10	32.60	3.31	27	38
Nurse 2	10	35.00	4.69	25	40
Nurse 3	10	33.70	4.79	24	38
Total	30	33.77	4.56	24	40
Patient Group H					
Nurse 1	10	18.80	2.49	16	24
Nurse 2	10	19.30	2.75	16	25
Nurse 3	10	19.60	3.95	15	26
Total	30	19.23	3.03	15	26
Patient Group D					
Nurse 1	10	17.70	1.64	15	20
Nurse 2	10	18.60	1.58	17	20
Nurse 3	10	17.80	1.48	15	20
Total	30	18.03	1.56	15	24
Nurse Raters					
1	30	23.03	7.20	15	38
2	30	24.30	8.18	17	40
3	30	23.70	7.92	15	38
Total	90	23.68	7.84	15	40

group has a total of 30 scores. Patient total classification scores range from 15 to 40, with an overall mean score of 23.68. Patient total classification scores among the three raters are about the same (mean scores are 23.03, 24.30, and 23.70). The mean patient total classification scores of the OPCS

follow a high-to-low sequence among patients in Group I (33.76), Group H (19.23), and Group D (18.03). Thus, the second research hypothesis is supported by the results.

To further test the differences of the mean patient total classification scores across the three patient groups for statistical significance, three statistical analyses were performed, a two-way ANOVA, a one-way ANOVA, and the Scheffé test. Statistical significance was set at $p \leq .05$ value.

Two-way ANOVA. Results from the two-way ANOVA are shown in Table 7. The interaction effect of nurse raters and patient groups on patient total classification scores was examined. The observed significance value was .868. Therefore, it appeared there was no interaction between the two variables. In the absence of an interaction, the test of main effects can be interpreted independently.

The F value associated with patient groups provided a test of the contrasted groups construct validity for the instrument. The F value (df = 2) associated with the patient groups was 223.825 (p <.001). Therefore, it appears that the difference among patient group means is extremely significant. At this point, the null hypothesis was rejected. On further investigation, the F value from two-way ANOVA was checked to determine whether rater differences affected patient scores. The F value associated with nurse raters was 1.172 (p =.315), indicating homogeneity among the three nurse raters. It was concluded that differences in nurse raters had no

Table 7

Two-Way ANOVA Associated With Patient Groups and Nurse Raters

Source of variation	Sum of squares	*df*	Mean square	*F*	Significance of *F*
Main effects	4626.044	4	1156.511	112.498	.000
patient group	4601.956	2	2300.978	223.825	.000
nurse raters	24.089	2	12.044	1.172	.315
2-way Interactions	12.911	4	3.228	.314	.868
patient group -nurses raters	12.911	4	3.228	.314	.868
Explained	4638.956	8	579.869	56.406	.000
Residual	832.700	81	10.280		
Total	5471.656	89	61.479		

effect on patient scores. This result further validated the findings for research hypothesis 1: "There will be significant interrater agreement on items of OPCS among nurse raters."

Because there was no interaction between nurse raters and patient groups, and the type of nurse raters had no effect on patient scores, patient group was deemed to be the only factor that affected patient scores. Thus, patient total classification scores for the three nurse raters were combined and again tested by patient group using statistics of one-way ANOVA and the Scheffé test.

One-way ANOVA and Scheffé test. Table 8 exhibits results from the one-way ANOVA on scores from the three patient groups. The F ($df = 2$, $p<.0001$) value of 230.1771 indicates that the difference among the group means was significant. However, results from the Scheffé test (see Table 9) partially support the contrasted groups construct validity; scores in Group I were significantly different from both Groups H and D, whereas no significant difference was demonstrated between scores in Groups H and D.

Table 8

One-Way ANOVA on Scores From the Three Patient Groups

Source of variation	*df*	Sum of squares	Mean square	*F* ratio	*F* prob.
Between groups	2	4601.9556	2300.9778	230.1771	.0000
Within groups	87	869.7000	9.9966		
Total	89	5471.6556			

In conclusion, results from both the patient-group mean scores and the ANOVAs statistical analyses supported research hypothesis 2: patient total classification scores differed in a high-to-low sequence among patients in medical-surgical intensive care units (Group I), patients not ready for discharge from medical-surgical units within 48 hours (Group H), and patients ready for discharge from medical-surgical units within 48 hours

Table 9

Scheffé Test on Scores From Three Groups of Patients

Group comparison	Mean difference	Significant *p* value
I vs. H	14.54	*(p <.05)
I vs. D	15.74	*(p <.05)
H vs. D	1.20	

Note:
I = patients in medical-surgical intensive care units.
H = patients not to be discharged in 48 hours from medical surgical units.
D = patients to be discharged in 48 hours from medical surgical units.
* = denotes pairs of groups significantly different at the .050 level when value actually compared between group mean is >2.2357.

(Group D), and demonstrated the construct validity of the instrument. However, the Scheffé post-hoc test revealed that the difference in patient scores between groups H and D did not reach a significant level. The statistically non-significant finding between Groups H and D shows that the subtle discrimination ability of the instrument is low.

Predictive Validity

To test the predictive validity of the instrument, patient total classification scores that were added from each of the 18 items were used. To test the predictive validity of the instrument, the relationship between patient total classification scores (independent variable) and raters'

estimation of nursing time (dependent variable) required for staffing was examined.

Research Hypothesis 3

The research hypothesis was, "Using the OPCS, there will be a positive correlation between patient total classification scores and the nursing time required by patients." The null hypothesis was, "Using the OPCS, there is no correlation between patient total classification scores and the nursing time that patients should receive."

Results From Statistical Analysis

A scatterplot showed an association between patient total classification scores and nursing time needed (Figure 3). Results from the scatterplot indicated that there was a curve pattern of association between the two variables; patient total score within the high score portion increased slower than the increases in staffing time required. An exponential curve provided the best fit for this plot. Nursing time scores were transformed into logarithmic form, and a second scatterplot was produced (Figure 4).

From Figure 4, a linear regression relationship between patient total classification scores and logarithmic nursing time was demonstrated. Increases in patient total classification scores linearly correlated with an increase in logarithm of nursing time/hours. Therefore, a Pearson's correlation was calculated to examine the predictive validity of the instrument. The correlation coefficient between patient total classification

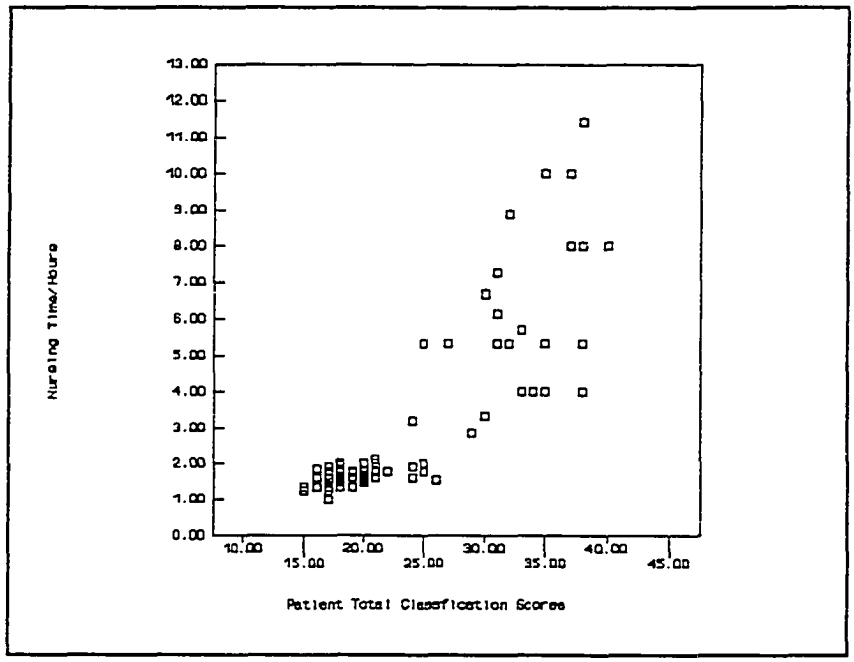


Figure 3. Relation between patient total classification scores and nursing time/hours needed.

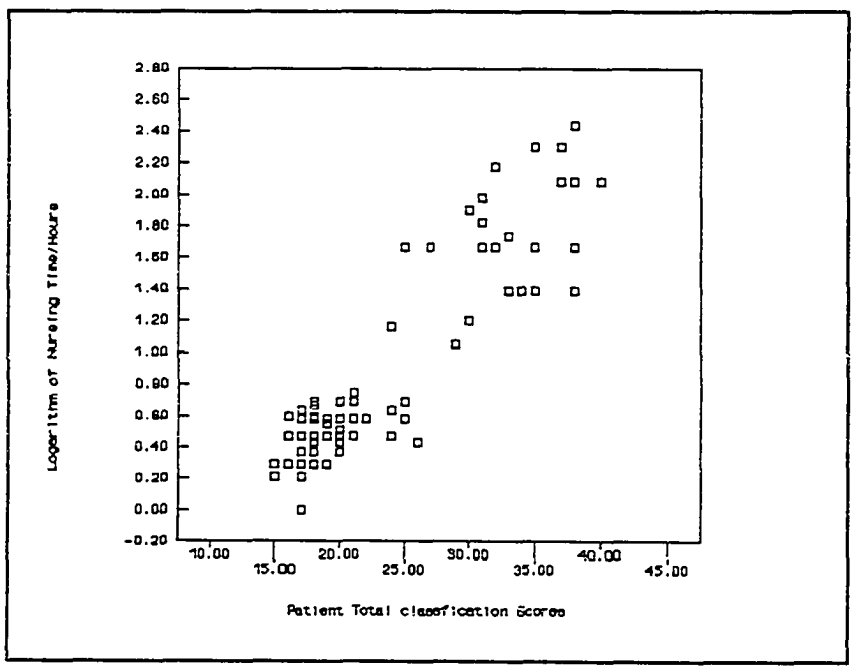


Figure 4. Regression of patient total classification scores on logarithm of nursing time/hours needed.

scores and the logarithmic nursing time/hours was .9249 (p <.001). The results rejected the null hypothesis and supported the research hypothesis 3 that there will be a positive correlation between patient total classification scores and the nursing time required by patients using the OPCS.

Calculation of Nursing Workload

The nursing care time estimated by the three nurse raters on a sample of 30 hospitalized patients ranged from 1 - 11.43 hours. Thus, five dissection points which divided the range of nursing time into six segments were made as 2, 4, 6, 8, and 10 hours. Calculated mean nursing time within each segment was 1.5 hours for a Class I patient; 2.8 hours for a Class II patient; and 4.5, 6.7, 8.2, and 10.5 hours for patients in Class III, IV, V, and VI, respectively.

Formulation of a Patient Classification System

A six-level PCS can be formulated through an estimation of five cut-off points within the range of patient total classification scores. As shown in Figure 4, there was a linear regression relationship between patient total classification scores and logarithmic nursing time. Therefore, these five cut-off points on patient total classification scores is determined by the five predetermined dissection points on the line of nursing time axis using a predictive equation. The predictive equation used to identify the five cut-off points on patient total classification scores from logarithmic nursing time was as follows:

$$Y = b_0 + b_n ln(t),$$

where variables in the equation were

Y = patient total classification scores (criterion),

b_0 = a constant (intercept value),

b_n = regression coefficient (slope),

ln = natural log (base e), and

t = nursing time (predictor).

From the linear regression procedure, a slope of 9.42 and a constant of 15.41 were generated in the column labeled B in the output, as shown in Table 10.

The five predetermined dissect points on nursing care time were 2, 4, 6, 8, and 10 hours. Based on the known slope and constant in the regression equation, the five cut-off points for patient total classification scores were measured as 21.6, 28.5, 32.3, 35.0, and 37.1. These five points were rounded to 22, 29, 32, 35, and 37. Using these five points as criteria, rules that divided all the patient total classification scores into six classes were established (see Table 11). Scores contained in each of the six classes (from low to high) ranged 14-21, 22-28, 29-31, 32-34, 35-36, 37-54. At this point, a one-tailed z test of the kappa value on patient classification levels of the instrument was calculated to check further the interrater reliability of the instrument at the patient classification level.

Table 10

Statistics for Variables in the Equation

Variable	B	SE_b	ß	T	Sig T
b_n (slope)	9.418560	.549922	.877059	17.127	.0000
(Constant)	15.405370	.626674		24.583	.0000

Table 11

Classification of Patient Total Classification Scores and Measurement of Nursing Workload

	Patient Class					
Categories	I	II	III	IV	V	VI
Patient score						
Minimum	14	22	29	32	35	37
Maximum	21	28	31	34	36	54
Mean nursing time/hours	1.5	2.8	4.5	6.7	8.2	10.5

Patient total classification scores were further transformed into levels of class according to the established classification rules.

A one-tailed *z* value was obtained by comparing kappa (K_a) value with the value of expected chance agreement (P_e) on patient classification levels. The critical value (1.65) was set at alpha = .05. The observed

agreement (P_o) on patient classification levels was 74% (see Table 12). Kappa value (.60) indicated a highly significant result when compared with the P_e (.36) using a one-tailed z test ($z = 6.31, p < .001$). This result provided extra support to research hypotheses 1 and further confirmed the reliability of the instrument.

Table 12

One-Tailed Z Test of Kappa Statistic on Patient Classification Levels of OPCS

Variables	P_o	P_e	K_a	SE_k	z
Levels of patient classification	.74	.36	.60	.10	6.31**

$N = 30$

Note: P_o = observed agreement; P_e = expected agreement; SE_k = standard error of kappa; K_a = kappa; $^{**}p < .001$

In conclusion, results from field studies had moderate high support to the research hypothesis 1, moderate low support to research hypothesis 2, and high support to research hypothesis 3. The OPCS had moderate interrater reliability, low contrasted groups construct validity and high predictive validity. Items 7 (sensory organ) and 15 (frequent tracheal suction) of the instrument were found not to be dependable. The difference of patient total classification scores between patient Groups H and D were not identified. This instrument needs revisions and a second

pilot study which is to be done in Taiwan before its implementation. In general, the reliability and validity demonstrated for OPCS were moderate and required further pilot study for patient classification.

Summary

This chapter presented results from the field pilot-testings of the instrument. Results are summarized in Table 13. The interrater reliability of the OPCS was supported by a one-tailed *z* test of kappa statistic on both the instrument's item level and the classification level. The contrasted groups construct validity of the instrument was supported by both one-way and two-way ANOVA. The predictive validity of the instrument was supported by a test of Pearson's correlation coefficient using the patient total classification scores and the logarithm of nursing time/hours. Because two items of the instrument were not dependable and Groups H and D were hard to differentiate, the overall instrument was considered to be moderately reliable and valid for an adult patient classification system to be used for medical-surgical patients in an acute-care teaching hospital.

Table 13

Summary of Analyses

Type of test	Null hypotheses (Ho) tested	Significant level	Ho rejected or not rejected
Interrater reliability	H1: There will be no significant interrater agreement on items of OPCS among the nurse raters	.05	Rejected
Contrasted groups construct validity	H2: There will be no difference in patient total classification scores among patients in medical-surgical intensive care units, patients not ready for discharge from the medical-surgical units within 48 hours, and patients ready for discharge from medical-surgical units within 48 hours	.001	Rejected
Predictive validity	H3: Using the OPCS, there will be no correlation between patient total classification scores and the nursing time required by patients	.05	Rejected

CHAPTER V
DISCUSSIONS, CONCLUSIONS, IMPLICATIONS AND RECOMMENDATIONS

In this study, a patient classification system was developed within the framework of Orem's self-care model. The newly developed instrument was named Orem's Patient Classification System (OPCS). A combination of descriptive, correlational, and factorial research design was applied to examine the instrument's reliability and validity. This chapter begins with a discussion of the findings as well as conclusions of the study. Subsequent sections address the implications of the study for nursing practice, education, research, and recommendations for further research.

Discussion

Discussions of the study include content related to issues of conceptual framework, methodology, as well as limitations of the study. The discussions are presented below.

Conceptual Issues

Because the Orem model is focused on an individual's self-care deficits (Meleis, 1991; Orem, 1991) in which nursing actions depend upon the nurse's ability to discern patient care needs, a PCS developed from Orem's model for the assessment of patient self-care deficits is deemed a

pertinent instrument for the determination of appropriate nurse staffing. The accomplishment of the study in the formulation and testing of the OPCS supported Fawcett's (1980; 1991) view that Orem's model has a conceptual framework that acts as a guide for research as well as for the administration of nursing service.

However, obstacles were encountered in the process of item formulation using the Orem model. These obstacles related to terminology used in the Orem framework of self-care requisites. As previously discussed in Chapter III, items in the framework of self-care requisites encompassed terms of similar concepts and terms of sets of concepts. Further, terms used in the framework were not easy to understand and translate into the patient-oriented state or traditional nursing action. When these terms were applied to items for the instrument, some terms had to be omitted, rewritten, reorganized by the researcher, or verified by content experts to maintain the essence of the framework while avoiding ambiguity among items. Through these instrument-development procedures, the study provided support to Hardy's (1974) views. Hardy (1974) proposed that much conceptual confusion occurs in theoretical areas from which nurses draw and that concepts must be reformulated through a process of relating the theoretical world to the empirical, arranging many items into a small number of classes (regrouping the bricks), and relating various concepts within a more general system of concepts. Consequently, refinement of the

concepts may contribute to a modification of the existing theory (Hardy, 1974).

Methodological Issues

Goodwin and Prescott (1981) maintained that an instrument should be assessed by more than one type of reliability. In the present study, the instrument's reliability was primarily assessed on the item-score level using a one-tailed *z* test of the kappa statistic. When the instrument was measured for contrasted groups construct validity, an additional measure testing for instrument reliability was generated on the patient total classification score level by the statistics of two-way ANOVA. After the instrument's classification system was established, an extra reliability test was conducted on the patient-class level using a one-tailed *z* test of the kappa statistic. Therefore, conclusions based on results of these three reliability tests are most convincing.

The one-tailed *z* test of the kappa statistic is a better measurement for testing the instrument's reliability rather than merely the use of the kappa statistic alone; the former is more sensitive in identifying the significance level of an item/instrument, while the latter only roughly identifies strength of agreement for the estimated item/instrument. Using data from Table 5 as an example, the kappa value for item 8 (orientation; K_a = .82) is greater than that of item 11 (health knowledge; K_a = .55). Because the expected agreement for item 8 (P_e = .83) is greater than that

of item 11 (P_e = .50), the one-tailed z test shows that the kappa value obtained from item 8 (K_a = .82; p <.01) is less significant than that of item 11 (K_a = .55; p <.001). If we only calculated kappa to test the instrument's reliability, we might misconclude that item 8 is more reliable than item 11. This finding questions Landis and Koch's (1977) interpretation to the value of kappa, such as the strength of agreement is perfect when kappa values range from .81 to 1.00, and the strength of agreement is moderate when kappa ranged from .41 to .60. Consequently, the study advises against using the kappa statistic as the sole measure in testing the instrument's reliability, for it may inaccurately represent instrument reliability.

If the reliability of the OPCS is measured by the percentage agreement, the overall observed agreement of the instrument was .87 (see Table 5) on the item-score level and dropped to .74 (see Table 13) on the patient-class level. Reasons for this conservative finding on the second approach may be explained partly by its use of a larger range of scores (Goodwin & Prescott, 1981); six categories (score from 1 to 6) were used for the second approach versus three (score from 1 to 3) for the first approach in the percentage agreement calculations.

Most reports in the literature recommended that the observed agreement for an instrument be over 90% (Giovannetti & Mayer, 1984; Haas, 1988; Huckabay & Skonieczny, 1981), few recommended 80% as the cut-off point (Shelley, 1984), and even fewer use 65% (Churness, Kleffel,

Onodera, & Jacobson, 1988). With more categories developed on instrument's patient-class levels in recent years, a single-standard recommendation as such is no longer appropriate for all PCSs. As the study shows, percentage agreement is partially dependent on the number of categories divided. When only a few categories, such as four (Helberg, 1989; Poulson, 1987), are involved in the calculation of the percentage agreement, a stringent standard would be appropriate, because differences among raters are reduced when fewer categories are available. When the categories are extended to five or six levels, an alternative standard must be established to accommodate the structural change of the instrument. It seems logical for this study to accept the reliability of a six-level PCS with an over 70% observed agreement on its patient-class level. In the study, the observed agreement of 74% for the instrument with six patient-class levels is therefore adequate.

A test of the instrument for the contrasted groups construct validity indicated that the patient total classification scores obtained from patients in the intensive care units (mean = 33.77) are greater than and statistically different (p <.05, see Table 9) from those of patients in Group H (mean = 19.23) as well as those patients in Group D (mean = 18.03). Scores obtained from Group H patients were higher than those from Group D. However, differences in scores for those two groups were not statistically significant at the alpha = .05 level. Thus, findings from this test only partly

support the second hypothesis, which states, "patient classification rating scores of the OPCS will demonstrate a high-to-low sequence among patients in medical-surgical ICUs, patients not ready for discharge from medical-surgical units within 48 hours, and patients ready for discharge from medical-surgical units within 48 hours."

It was found in the literature that the prospective payment system had shortened the average length of patient stay in acute care facilities (Langer, Drinka, & Voeks, 1991). Those discharged early are cared for in the home (Goulart, 1991) and benefit from cost savings as well as continuity of family routine (Rucker & Harrison, 1974), and are becoming more similar in acuity of illness with those patients who are not within 48 hours of discharge. The growing demand for nursing care and medical treatment for patients discharged early has contributed significantly to the development of home health agency nurses (Churness, Kleffel, Onodera, & Jacobson, 1988).

A change in the traditional treatment regimens is also reflected in the data collection processes of this study. While collecting data, the researcher found that some acutely ill patients were included in Group D. For example, an overweight female patient classified as ready for discharge was found to have an indwelling intravenous line as well as a full-leg cast. Once at her home, she would be receiving intravenous antibiotic therapy administered by the visiting nurses. When these acutely ill patients were

selected in Group D, the acuity level between patients in the Groups H and D tended to be similar.

Reports from the literature and findings from the data collection procedure reflected the accuracy in the findings from the contrasted groups construct validity test, which indicated no support for the difference in scores between Group H (mean score = 19.23) and Group D patients (mean score = 18.03). The influence of the prospective payment system makes the fourth assumption, upon which the latter portion of the hypothesis is established, to be unwarranted and the subdivision of patients in the medical-surgical units unnecessary. The fourth assumption stated that "the nursing hours required for patients who are not ready for discharge from the hospital within 48 hours are greater than for those ready for discharge from the hospital within 48 hours." Based on the previous findings, it can be concluded that because of the influence of the prospective payment system and shorter lengths of stay in the hospital, the discrimination abilility of the OPCS was low. In other words, the subtle differences in patients will not be found with the OPCS, unless precise points are collapsed into broader categories.

According to types of educational preparation, there were several types of nurses, such as Registered Nurses (RNs), providing care to patients as well as to healthy individuals. According to the hospital policy in the current study setting, the study only involved RNs to classify patients.

Future studies will not specify the required level of preparation for nurses classifying the patients because the instrument is to be used in a country where nurses' tasks have not yet been specified by the types of professional nursing preparation.

Limitations of the Study

In some cultures, patient's family members may share in the nursing care activities and therefore influence nursing workload and significantly affect health care costs. At present, this study has not yet considered the cultural influence in nursing workload measurement, for it would make the system too complicated during the initial period of instrument development.

This study was limited to medical-surgical patients between 18 and 65 years of age from one teaching hospital. Thus, the findings may not be generalizable to all patients in all types of hospitals and to all health care systems. Finally, the size of the sample was a limitation. Only 30 patients were classified. Developing a six-level PCS based on information collected from 30 hospitalized patients may not provide all the information necessary for classification.

Conclusions

The purpose of this study was to develop a six-level acuity-based factor-type medical-surgical PCS derived from Orem's (1991) nursing model and to test its reliability and validity in an acute-care teaching hospital in the southeast region of the United States. A future goal is to further

validate and implement this OPCS in medical-surgical units of an acute-care teaching hospital in Taiwan. The content validity of the instrument has been supported by the content experts using the CVI measurement. In this section, the conclusions of the study are presented in association with Orem's model and the findings.

Orem's Model

As in the previous discussion, the OPCS was formulated based upon the framework from Orem's model, thus, indicating that Orem's model has a conceptual framework that acts as a guide for research as well as for administration of nursing service (Fawcett, 1980; 1991). Obstacles encountered while formulating items of the instrument were related to Orem's use of terminologies in the framework of self-care requisites. Some terms used in the framework are ambiguous; others are not easily understood. Through a refinement and transformation procedure, these terms were converted into day-to-day nursing practices, and the first draft of the instrument was developed.

Findings

Three research hypotheses were formulated to test the instrument for interrater reliability, contrasted groups construct validity, and predictive validity. Based upon findings from the research hypotheses, conclusions for the research questions were made.

Interrater Reliability

The first research hypothesis, which states, "there will be a significant interrater agreement on items of the OPCS among the nurse raters," was supported by the one-tailed z test of the kappa statistic ($z = 40.97, p < .001$, see Table 5). When the instrument was examined item-by-item for the interrater reliability, 16 out of the 18 items exhibited a significant agreement ($z = 2.67 - 11.37, p < .01$, see Table 5). The significance of raters' observed agreement did not occur in items 7 (sensory organ) and 15 (tracheal suction). Thus, it was concluded that the overall instrument was considered to be a reliable one and that two of the items required further revisions thereby making all items of the instrument equally reliable.

In addition, the two-way ANOVA test revealed that there was no significant difference among the three nurse raters on patient total classification scores ($F = 1.172, p = .315$, see Table 7). The one-tailed z test of the kappa statistic also demonstrated a statistically significant interrater agreement among the raters on patient classes ($z = 6.31, p < .001$, see Table 13). Both findings were consistent with findings from the observed agreement on item scores of the instrument, therefore, the interrater reliability of the instrument was further supported.

Contrasted Groups Construct Validity

The second research hypothesis states that "patient classification rating scores of the OPCS will demonstrate a high to low sequence among

patients in medical-surgical ICUs, patients not ready for discharge from medical-surgical units within 48 hours, and patients ready for discharge from medical-surgical units within 48 hours." The mean total classification scores among the three patient groups do follow the proposed sequence: 33.77 in Group I, 19.23 in Group H, and 18.03 in Group D.

The difference in the mean patient total classification scores among the three groups reached a statistically significant level when the scores were tested using both two-way ANOVA ($F = 223.825, p < .001$, see Table 7) and one-way ANOVA (230.177, $p < .0001$, see Table 8). When the patient total classification scores among the three groups were examined by the Scheffé post-hoc test, scores in Group I were statistically different from scores in Groups H and D ($p < .05$), but the difference between scores in Groups H and D were not statistically significant at alpha = .05 level. As in the previous discussion, the contrasted groups construct validity of the instrument is sufficient when patient total classification scores obtained in Group I were greater than and statistically different from those of patients in Groups H and D. Because test findings supported the first part of the research hypothesis, the OPCS is valid in differentiating patients in different acuity units.

Predictive Validity

The third research hypothesis, which states, "there will be a positive correlation between patient total classification rating scores and the nursing

care time required by patients using the OPCS," was supported by the Pearson's correlation coefficient (r =.9249, p <.001). Thus, it can be concluded that the OPCS was valid in predicting nursing time required by patients being classified.

The research question applied to the study is, "What is the support for reliability and validity of a patient classification instrument that is derived from Orem's model for use in an acute care clinical setting?" Based on findings of the study, answers to this research question include:

1. the overall instrument is reliable for use in patient classification,
2. the two items with low interrater reliability require revisions in order to make all items of the instrument equally reliable,
3. the OPCS is valid in differentiating patients in different acuity units, and
4. the OPCS is valid in predicting nursing time required by the patients being classified.

In summary, the OPCS was developed from Orem's model and is a reliable and valid instrument for classification of patients in an acute-care teaching hospital.

Implications

Implications are discussed in terms of nursing practice, nursing research, and nursing education. These discussions are presented below.

Nursing Practice

Integration of the Orem nursing model with a PCS can facilitate nursing assessment in clinical settings. With a nursing framework contained in the OPCS, this instrument may be adapted easily for use in patient assessment. Thus, the OPCS is not only for use in patient classification but also contributes to nursing assessment.

Because the OPCS was developed based on the identification of patient self-care deficits, it is able to assess the number of nurses required for a group of patients. Other subsequent systems, such as nursing cost or budgeting and billing systems easily can be established to assist nurse administrators in managing their department cost-effectively.

Nursing Research

Findings from this study show that the reliability of the OPCS measured by the observed agreement is partially dependent on the number of categories developed. Nurse researchers can utilize this information to establish standards for reliability testing of a PCS. As in a previously mentioned example, when a six-level PCS is developed, a minimum of 70% observed agreement on patient-class may be applied in the testing of the instrument's reliability.

Findings of the study also indicate that the kappa statistic alone is not a dependable method for testing the reliability of an instrument. Instead, the one-tailed *z* test of kappa statistic is a better choice. Nurses

can utilize this information in testing, revising, and retesting an instrument accordingly.

Nursing Education

Through the development of a PCS, knowledge that expands our understanding of the use of Orem's nursing model has been generated. Nurse educators may include this information in professional nursing education programs to extend knowledge of nursing theory.

Findings from this study also show that the reliability of the OPCS measured by the observed agreement is partially dependent on the numbers of categories divided. Nurse educators may utilize this information to establish standards for the PCS training program.

Recommendations for Further Research

The significant as well as non-significant results of findings from the study suggest areas for future discussion and testing. In future studies, indicators having low interrater reliability (item 7, sensory organ, and item 15, tracheal suctioning) need to be refined and re-tested to make certain that all items of the instrument are reliable. It is also anticipated that patients selected for future study need only be divided into two groups (Groups I and H) to test the contrasted groups construct validity.

It is important to note that the reliability and validity generated for the OPCS are based on one selected sample of adult medical-surgical patients. Currently, there is no evidence that testing with other populations

would generate similar results. A replication of the study using adolescent or elderly patients may be more representative of the population, and might reveal even greater discrepancies in the consensus measures.

Taking cultural influences into consideration, in future studies, assistance from family members might be calculated by creating an additional category for each item. This category would take into account care provided by family members and would exclude the equivalent amount of family work from the total measurement of the nursing workload.

Since nursing time estimated by the nurse raters was based on an 8-hour basis, in the future, it would require that nursing workload of evening and night shifts also be calculated. In addition, when there is sufficient support for conducting a future study, the required nursing time measured by a time-and-motion study or a video study may prove worthwhile for use in order to conduct a check of the predictive validity of the instrument.

Further research in revision and testing of the instrument remains an important part in the development of the body of nursing knowledge and the determination of the usability of the instrument in nursing practice as well as in nursing administration.

Summary

This research has yielded several important contributions to the discipline of nursing. Through the development of a PCS, knowledge that expands our understanding of the use of Orem's nursing model has been

generated. An empirical test of the conceptualizations of Orem's model has been done. Results from the three research hypotheses provide a positive answer to the research question that a PCS derived from Orem's model is reliable and valid for use in an acute-care teaching hospital. Implications of the study for nursing practice, as well as nursing research, and education were demonstrated. Recommendations for further research related to replication of the study were suggested.

REFERENCES

Abdellah, F. G., & Levine, E. (1979). Better patient care through nursing research (2nd ed.). New York: Macmillan.

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

Ambutas, S. (1987). Evaluating a patient classification system. Dimensions of Critical Care Nursing, 6, 364-367.

Ambutas, S. (1988). A comparison of two patient classification systems for an MICU. Nursing Management, 19(9), 64A-64H.

Arenth, L. M. (1985). The development and validation of an oncology patient classification system. Oncology Nursing Forum, 12(6), 17-22.

Aydelotte, M. K. (1973). Nurse staffing methodology - A review and critique of selected literature (DHEW Publication No. NIH 73-434). Bethesda, MD: National Institute of Health.

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

Bermas, N. F., & Slyck, A. V. (1984). Patient classification systems and the nursing department. Hospitals, 58(22), 99-100.

Budd, M. C., & Propotnik, T. (1989). A computerized system for staffing, billing and productivity measurement. Journal of Nursing Administration, 19(7), 17-22.

Burns, N., & Grove, S. K. (1987). The practice of nursing research: Conduct, critique and utilization. Philadelphia: W. B. Saunders.

Catanzaro, W. (1988). Nursing research: Theory and practice. St. Louis: C. V. Mosby.

Chagnon, M., Audette, L. M., Lebrun, L., & Tilquin, C. (1978). A patient classification system by level of nursing care requirements. Nursing Research, 27(2), 107-113.

Churness, V. H., Kleffel, D., Onodera, M. L., & Jacobson, J. J. (1988). Reliability and validity testing of a home health patient classification system. Public Health Nursing, 5(3), 135-139.

Claussen, E. (1955). Categorization of patients according to nursing care needs. Military Medicine, 116(3), 209-214.

Cottey, K., Nauert, L., & Wills, S. (1992). Modifying a hospital-wide PCS for burn wound care. Nursing Management, 23(4), 58-62.

Croft, A. (1993). A psychiatric patient classification system that works! Nursing Management, 24(11), 66-70, 72.

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

Davis, V. J., & Bertram, D. L. (1991). Staffing methodologies. In D. L. Bertram & J. L. Wilson (Eds.), Financial management in critical care nursing (pp. 131-143). Baltimore: Williams & Wilkins.

De Groot, H. A. (1989a). Patient classification system evaluation: Part 1: Essential system elements. Journal of Nursing Administration, 19(6), 30-35.

De Groot, H. A. (1989b). Patient classification system evaluation: Part 2: System selection and implementation. Journal of Nursing Administration, 19(7), 24-30.

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.

Dijkers, M., Paradise, T., & Maxwell, M. (1986). Pitfalls of using patient classification systems for costing nursing care. In F. A. Shaffer (Ed.), Patients and purse strings: Patient classification and cost management, New York: National League for Nursing.

Donnelly, L. J. (1981). Patient classification: An effective management tool. Nursing Management, 12(11), 42-43.

Dowman, R., Spence, A., & Sankaran, S. (1990). Patient classification. New Zealand Nursing Journal, 83(5), 20-22.

Ebener, M. (1985). Reliability and validity basics for evaluating classification systems. Nursing Economics, 6, 324-327.

Erikson, E. H. (1963). Childhood and society (2nd ed.). New York: W. W. Norton.

Fawcett, J. (1980). A framework for analysis and evaluation of conceptual models of nursing. Nurse Educator, 5(6),10-14.

Fawcett, J. (1991). Analysis and evaluation of conceptual models of nursing (2nd ed.). Philadelphia: F. A. Davis.

Fetter, R. B., Brand, D. A., & Gamache, D. G. (Eds.). (1991). DRGs: Their design and development. Ann Arbor, MI: Health Administration Press.

Fetter, R. B., Mills, R. E., Riedel, D. C., & Thompson, J. D. (1977). The application of diagnostic specific cost profiles to cost and reimbursement control in hospitals. Journal of Medical Systems, 1(2), 137-149.

Fleiss, J. I. (1981). Statistical methods for rates and proportions (2nd ed.). New York: John Wiley & Sons.

Fox, D. (1982). Fundamentals of research in nursing (4th ed.). Norwalk, CT: Appleton-Century-Crofts.

Fries, B. E., & Cooney, L. M. (1984). A patient classification system for long-term care. (DHHS Publication No. HCFA 84-03179). Baltimore, MD: Office of Research and Demonstration.

Gender, A. R. (1989). Developing and computerizing a patient classification system in a rehabilitation nursing setting. Rehabilitation Nursing, 14(2), 58-63, 72.

Giovannetti, P. (1978). Patient classification systems in nursing: A description and analysis. (DHEW Publication No. (HRA) 78-22, HRP-0500501). Hyattsville, MD: U.S. Department of Health, Education, and Welfare.

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

Giovannetti, P., & Johnson, J. M. (1990). A new generation patient classification system. Journal of Nursing Administration, 20(5), 33-40.

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

Goodwin, L. D., & Prescott, P. A. (1981). Issues and approaches to estimating interrater reliability in nursing research. Research in Nursing and Health, 4, 323-327.

Goulart, D. T. (1991). Educating the cardiac surgery patient and family. Journal of Cardiovascular Nursing, 3(3): 1-9.

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

Hardy, M. E. (1974). Theories: Components, development, evaluation. Nursing Research. 23(2), 100-107.

Hay, M. A., & Nelson, L. M. (1988). Client classification: A "needs" approach. Journal of Psychosocial Nursing & Mental Health Services, 26 (12), 23-30.

Hastings, C. E. (1987). Classification issues in ambulatory care nursing. The Journal of Ambulatory Care Management, 10(3), 50-64.

Helberg, J. L. (1989). Reliability of the nursing classification index for home healthcare. Nursing Management, 20(3), 48-56.

Helmer, F. T., & Halbert, N. (1987). The development of a patient classification for pediatric orthopaedic units: A case study. Orthopaedic Nursing, 6(6), 32-38.

Hoffman, F. (1988). Nursing productivity assessment and costing out nursing services. Philadelphia: J. B. Lippincott.

Hoffman, F., & Wakefield, D. S. (1986). Ambulatory care patient classification. Journal of Nursing Administration, 16(4), 23-30.

Huckabay, L. M. (1981). Patient classification: A basis for staff. New York: NLN.

Huckabay, L. M., & Skonieczny, R. (1981). Patient classification systems: The problems faced. Nursing and Health Care, 2, 89-102.

International classification of diseases, 9th revision, clinical modification, ICD-9-CM. (1978). Ann Arbor, MI: Commission on Professional and Hospital Activities.

Jackson, B. S., & Resnick, J. (1982). Comparing classification systems. Nursing Management, 13(11), 13-19.

Jelinek, R. (1989). Patient classification systems. In D. A. Gillies (Ed.), Nursing management: A systems approach (pp.288-305). Philadelphia: W. B. Saunders.

Jennings, B. M., Rea, R. E., Antopol, B. B., & Carty, J. L. (1989). Selecting, implementing, and evaluating patient classification systems: A measure of productivity. Nursing Administration Quarterly, 14(1), 24-35.

Johnson, J. M. (1989). Quantifying an ambulatory care patient classification instrument. Journal of Nursing Administration, 19(11), 36-42.

Joint Commission on Accreditation of Healthcare Organizations. (1982). Accreditation manual for hospitals. Chicago: Author.

Joint Commission on Accreditation of Healthcare Organizations. (1992). Accreditation manual for hospitals. Chicago: Author.

Katz, S., & Akpom, C. A. (1976). Twelve indexes of ADL. Medical 14(Suppl. 5), 116-118.

Kessler, D. M., Kessler, D. H., & Knibloe, M. E. (1990). A dialysis patient classification system for establishing nurse/patient ratios. Journal of American Nephrology Nurses' Association, 17(5), 367-371.

King, I. M. (1975). Patient aspects. In L. J. Shuman, R. D. Speas, & J. P. Young (Eds.), Operations research in health care: A critical analysis (pp. 3-20). Baltimore: Johns Hopkins University Press.

Kinley, J., & Cronenwett, L. R. (1987). Multiple shift patient classification: Is it necessary? Journal of Nursing Administration, 17(2), 22-25.

Kyle, B. A., & Kinder, S. (1990). Costing out nursing: Combining PCSs, DRGs, and standards of care. Health Care Supervisor, 8(2), 46-56.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. Biometrics, 33, 159-174.

Langer, E., Drinka, P. J., & Voeks, S. (1991). Readmissions and acuity in the nursing home: How will the nursing homes manage? Journal of Gerontological Nursing. 17(7): 15-19.

Leatt, P., Bay, K., & Stinson, S. M. (1981). An instrument for assessing and classifying patients by type of care. Nursing Research, 30(3), 145-150.

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

Lovett, R. B., Wagner, L., & McMillan, S. (1991). Validity and reliability of a pediatric hematology oncology patient acuity tool. Journal of Pediatric Oncology Nursing, 8(3), 122-130.

Marks, R. J., Simons, R. S., Blizzard, R. A., & Browne, D. R. G. (1991). Predicting outcome in intensive therapy units: A comparison of APACHE II with subjective assessments. Intensive Care Medicine, 17, 159-163.

Mateo, M. A., & Kirchhoff, K. T. (1991). Conducting and using nursing research in the clinical setting. Baltimore: Williams & Wilkins.

McKenzie, D. A. (1991). A proposed prototype for identifying and correcting sources of measurement error in classification systems. Medical Care, 29(6), 521-530.

McLaughlin, F. E., & Marascuilo, L. A. (1990). Advanced nursing and health care research. Philadelphia: W. B. Saunders.

Medicus Systems Corporation. (1990). Staffing/productivity system: Type VI methodology. Evanston, IL: Author.

Meleis, A. I. (1991). Theoretical nursing: Development & progress (2nd. ed.). Philadelphia: J. B. Lippincott.

Meyer, D. (1978). Work load management system ensures stable nurse-patient ratio. Hospitals, 52(5), 81-85.

Mowry, M. M., & Korpman, R. A. (1985). Do DRG reimbursement rates reflect nursing costs? Journal of Nursing Administration. 15(7,8), 28-34.

Mowry, M. M., & Korpman, R. A. (1986). Managing health care cost, quality and technology. Rockville, MD: Aspen.

Nagaprasanna, B. (1988). Patient classification systems: Strategies for the 1990s. Nursing Management, 19(3), 105-112.

Niemeier, D. F., & Reed, G. B. (1985). A monitoring system for patient classification in critical care. Dimensions of Critical Care Nursing, 4(2), 110-118.

Nunnally, J. C. (1978). Psychometric theory (2nd ed.). New York: McGraw-Hill.

O'Leary, C. (1991). A psychiatric patient classification system. Nursing Management, 22(9): 66.

O'Neil, S. (1972). The application and methodological implications of behavior modification in nursing research. In M. Batey (Ed.), Communicating nursing research. Boulder, CO: Western Interstate Commission for Higher Education.

Orem, D. E. (1971). Nursing: Concepts of practice. New York: McGraw-Hill.

Orem, D. E. (1985). Nursing: Concepts of practice (3rd ed.). New York: McGraw-Hill.

Orem, D. E. (1991). Nursing: Concepts of practice (4th ed.). St. Louis: Mosby.

Pardue, S. F., & Dick, C. T. (1986). Patient classification: Illness acuity and nursing care needs. Journal of Psychosocial Nursing, 24(12), 23-30.

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

Polit, D. F., & Hungler, B. P. (1991). Nursing research: Principles and methods (4th ed.). Philadelphia: J. B. Lippincott.

Poulson, E. (1987). A method for training and checking interrater agreement for a patient classification study. Nursing Management, 18(9), 72-80.

Roehrl, P. K. (1979). Patient classification: A pilot test. Supervisor Nurse, 10(2), 21-27.

Rosko, M. D. (1988). DRGs and severity of illness measures: An analysis of patient classification systems. Journal of Medical Systems, 12(4), 257-274.

Rucker, R. W., & Harrison, G. W. (1974). Outpatient intravenous medications in the management of cystic fibrosis. Pediatrics, 54, 358-360.

Ruman, J. M., & Nelson, D. L. (1987). Patient classification systems capture nursing costs. Healthcare Financial Management, 39(12), 84-87.

Schroder, P. J., Washington, W. P., Deering, C. D., & Coyne, L. (1986). Testing validity and reliability in a psychiatric patient classification system. Nursing Management, 17(1), 49-54.

Schwamb, J. (1989). A maternity patient classification system. Nursing Management, 20(11), 66-71.

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

Sherman, J. J. (1990). Costing nursing care: A review. Nursing Administration Quarterly, 14(3), 11-17.

Slyck, A. V. (1991). A systems approach to management of nursing services - Part II: Patient classification system. Nursing Management, 22(4), 23-25.

Soeken, K. L., & Prescott, P. A. (1986). Issues in the use of kappa to estimate reliability. Medical Care, 24(8), 733-741.

Stepura, B. A., & Miller, K. (1989). Converting nursing care cost to revenue. Journal of Nursing Administration, 19(5), 18-22.

Strom, C. A. (1993). Combining patient classification and nursing documentation. Nursing Management, 24(2), 79.

Suen, H. K., & Ary, D. (1989). Analyzing quantitative behavioral observational data. Hillsdale, NJ: Lawrence Erlbaum.

Sullivan, E. J., & Decker, P. J. (1985). Effective management in nursing, Menlo Park, CA: Addison-Wesley.

Swansburg, R. C., & Sowell, R. L. (1992). A model for costing and pricing nursing service. Nursing Management, 23(2), 33-36.

Thomas, B. S. (1990). Nursing research: An experiential approach. St. Louis: C. V. Mosby.

Thompson, J. D., & Diers, D. (1991). Nursing resources. In R. B. Fetter, D. A. Brand, & D. Gamache (Eds.). DRGs: Their design and development (pp. 121-183). Ann Arbor, MI: Health Administration Press.

Van Der Walt, A. M. (1992). Patient classification in the Groute Schuur Hospital Region, Republic of South Africa. Nursing Administration Quaterly, 16(2), 43-46.

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

Verran, J. A. (1986). Testing a classification instrument for the ambulatory care setting. Research for Nursing and Health, 9, 279-287.

Wagner, D. P., & Draper, E. A. (1984). Acute physiology and chronic health evaluation (APACHE II). Health Care Financing Review, 6(Annal. Suppl.), 91-105.

Wake, M. M. (1990). Nursing care delivery systems: Status and vision. Journal of Nursing Administration, 20(5), 47-51.

Waltz, C., & Bausell, R. B. (1986). Nursing research: Design, statistics, and computer analysis. Philadelphia: F. A. Davis.

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

Washington, C. C., & Moss, M. (1988). Pragmatic aspects of establishing interrater reliability in research. Nursing Research, 37(3), 190-191.

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

Wolfe, H., & Young, J. P. (1965). Staffing the nursing unit - Part II: The multiple assignment technique. Nursing Research, 14(4), 299-303.

Woods, N. F., & Catanzaro, M. (1988). Nursing research: Theory and practice. St. Louis: C. V. Mosby.

Woolson, R. F. (1987). Statistical methods for the analysis of biomedical data. New York: John Wiley & Sons.

Wright, M. (1954). Improvement of patient care. New York: G. P. Putnam & Sons.

Zembala, S. A. (1993). Managing the costs of nursing care delivery. Nursing Administration Quarterly, 17(4), 74-79.

APPENDIX A

INSTRUMENT AS PRESENTED FOR A FIRST ROUND EVALUATION OF CONTENT VALIDITY

Directions for Patient Classification

This instrument will permit you to score an adult medical-surgical patient on 15 different patient self-care requisites/deficits and 4 optional patient self-care deficits. Each item consists of groups of indicators that are further classified into three modes of nursing care complexity. Attached are guidelines to Orem's patient classification system, a listing of Orem's classification indicators, and definitions of the indicators for your use.

Please enter the patient's demographic data in the area provided. Fill- in patient's major diagnosis, age, sex, and race accordingly. Then, estimate the maximum number of patients you can take care of per shift if all the patients were like the one you are classifying. Round one position after decimal point. Finally, score the patient's requisites/deficits for each item in the following:

Minimum nursing care need = 1

Moderate nursing care need = 2

Intensive nursing care need = 3

* Remember, for each item, the score should be either a 1, 2, or 3 on each of the 15 items.

Guidelines for Using the Orem's
Patient Classification System (OPCS)

1. All hospitalized patients should be classified daily by their R.N. caregivers no later than 10:00 a.m.

2. Classification should include **all** patients:

   a. Present on the unit as of 10 a.m. who have not been discharged.

   b. Who have been discharged but are still present on the unit.

   c. Off the unit for surgery, treatments, examinations or are on passes and will return by 6:00 p.m.

3. Do **not** classify patients:

   a. Discharged or transferred prior to 10:00 a.m.

   b. In surgery and will not return to the unit prior to 06:00 p.m.

   c. Off the unit for examinations, treatments or are on passes but will not return prior to 6:00 p.m.

4. Complete Demographic Data on the Patient Classification Form (PCF).

5. Obtain data from observation of the patient, the patient's charts and the shift-change nursing report.

6. Classify patient according to the Listing of Orem's Patient Classification Indicators.

7. Record score for each item on the PCF.

8. Enter the total patient classification score on the PCF.

9. Initial the PCF and then give it to the registered nurse on the unit who is responsible for data gathering.

Orem's Patient Classification Indicators

Item	Nursing Care Needed		
	Minimum (Score=1)	Moderate (Score=2)	Intensive (Score=3)
1. Air intake	No respiratory difficulty	Some difficulty in breathing /Natural drainage of chest tube /Old tracheostomy	Application of machine
2. Fluid intake	Regular fluid intake	Fluid intake restriction /Single regular IV	Single critical IV/Telemetry or EKG monitoring continuously
3. Food intake	Regular or soft diet	NPO/Modified diet intake /Nausea or vomiting	Tube feeding /N-G tube with suction machine
4. Bladder elimination	Regular elimination	Difficulty in urination /Use of catheter	Urinary retention /Incontinence
5. Bowel elimination	Proficient use of colostomy by self /Regular bowel movements	Use of colostomy /Difficulty in bowel movements /Tube to abdomen /Incontinent	New colostomy /Active G-I hemorrhage
6. Hazard prevention	No sensory impairment or deficit /Alert	Some sensory impairments	Extensive sensory deficits

7. Activity	Ambulatory /Adequate range of motion	Weak motor strength /Local restraints /Use of prosthetics	Absolute bed rest /Completely immobilized motor activity
8. Rest (comfort)	Normal rest or sleep pattern /Unconscious	Transfer in or out /Minor anxiousness, restlessness, or in moderate pain	Admitted or discharged /Major anxiousness, restlessness, or in severe pain
9. Promotion of normalcy	Compliant	Intermittently Compliant	Non-compliant
10. Health knowledge	Not ready for education /Minor knowledge deficit	Scheduled education /Moderate knowledge deficit & ready for education	Pre-operative education/Major knowledge deficit & ready for education
11.Social support (SS)	Family and friends available	Limited SS	Stressful family situation
12. Financial security	Has health insurance for hosp or supported by research funding	Has social service aids /No insurance and no financial problem	No insurance and has financial problem
13. Lab data collection	Minor bedside data collection	Moderate bedside data collection	Complicated bedside data collection
14. Medical regimen	Oral medications	IV medications ≤2 doses/shift	IV medications ≥3 doses/shift
15. Surgical regimen	No Surgery /Post-op ≥48 hours/No change dressing (CD) or CD qd	Post-op ≥24 hours/Drainage tube to wound other than chest & abdomen/CD bid	Pre-op physical preparation/Post-op <24 hours/CD ≥tid

Optional Classification Items:

Optional Item	Nursing Care Complexity		
	Minimum (Score = 1)	Moderate (Score = 2)	Intensive (Score = 3)
16. Frequent tracheal or endo-tracheal suctioning	Suctioning q4h	Suctioning q2h	Suctioning q1h or more/On CPR
17. Multiple IV lines	Two IV lines	Three IV lines	Four IV lines or more
18. Vital sign (VS) or intake and output (I&O)	Q1h, Q30' & Q15' for VS or I&O ≥2 hours	Q30' & Q15' for VS or I&O ≥2 hours	Q15' for VS or I&O ≥2 hours
19. Special examinations or treatment procedures	Bedside exam or treatment <1 hour /Exam or treatment at other department	Bedside complicated exam or treatment ≥1 hour, but <2 hours	Bedside complicated exam or treatment ≥2 hours

NOT FOR USE BY DATA COLLECTOR:

1) Total score is the sum of all scores obtained from each item in the columns.

2) Minimum score is 15 and maximum score is 57.

Definitions of Patient Classification Indicators

1a. No respiratory difficulty.
1. Normal respiration.
2. Stoma in trachea.

1b. Some difficulty in breathing.
1. Shortness of breath.
2. O_2 prn use.
3. Productive cough.
4. Frequent non-productive cough.

1c. Old tracheostomy.
1. Intermittent use of cap on the site of tracheostomy.
2. Use of tracheostomy but no suction needed.

1d. Application of machine.
1. Use of suction machine for chest tube, tracheostomy.
2. The application of ventilator.

2a. Fluid intake restriction.
1. Limit or force fluids.

2b. Single regular IV.
1. Single IV with plain fluids or drugs such as insulin, bicarbonate.
2. Heparin lock IV.

2c. Single critical IV.
1. IV with blood transfusion or life-saving drugs in the bottle such as anti-hypotension/hypertension or chemotherapy agents.
2. IV infusions requiring special catheter to heart.
3. Special designed IV line to measure venous pressure or cranial pressure (e.g., CVP line, ICP line).

3a. Modified diet.
1. Special diet such as clear liquid, high caloric, sodium restriction, diabetic diet.

3b. Tube feeding.
1. Use of N-G tube, G tube, or TPN tube.

4a. Difficulty in urination.
1. Complaint of difficulty, frequency, burning, and/or urgency.

4b. Use of catheter.
1. Use of condom, or intermittent/indwelling catheter.

5a. Difficulty in bowel movements.
1. Abdominal distention, constipation, fecal impaction, diarrhea.

5b. Use of colostomy.
1. Colostomy made more than 3 days and not proficient in use by self.

5c. New colostomy.
1. Have colostomy <3 days.

5d. Active G-I tract hemorrhage.
1. Vomiting or bleeding from anus >50 ml within 24 hours of the classification time.

6a. Some sensory impairments.
1. Impaired hearing (use of hearing aid) or seeing (e.g., snowflake cataract, excluding use of contact lenses or eye glasses).
2. Vertigo.
3. Confused.
3. Delayed development.

6b. Extensive sensory impairments.
1. Deaf, blind or almost blind (can see only lights and no figures), mute or unconscious.

7a. Ambulatory.
1. Walks freely, proficient use of aids for motor activity (e.g., crutches; artificial limbs).

7b. Weak motor strength.
1. Requires side-rails.
2. Bed rest with bathroom privilege.
3. Up with assistance such as for general weakness in parts of body or legs (e.g., post-operation days, complications of diabetic neuropathy or congenital heart failure).
4. Impaired range of motion (stiffness on parts of body) such as PP cast, cranio-fixation, hemiplegia or paraplegia.
5. New prosthesis not used proficiently.

7c. Local restraints.
1. Use of extremity or vest restraints.

7d. Completely immobilized or on restraints.
1. Cranio-traction.
2. Physically unresponsive or only responsive to pain and pressure (e.g., unconscious, quadriplegia).
3. Burns >40%.

8a. Transfer in or out.
1. Transferred into the unit before 10 a.m. or out of the unit after 10 a.m. of the classification date.

8b. Minor anxiousness, restlessness, or in moderate pain.
1. Expresses anxiety and is restless but not agitated (e.g., fears loss of job, fears the coming special treatment).
2. Has one or two special therapies or interventions (e.g., spinal cord puncture).
3. Sleepless or insomnia.
4. Speaks foreign language and some local language.
5. Speaks only foreign language and interpreter available.
6. Complaint of intermittent pain or discomfort and/or needs some pain medicine.

8c. Admitted or discharged.
1. Arrived in unit before 10 a.m. of the classification date.
2. Scheduled to be discharged from the unit but still present after 10 a.m. of the classification date.

8d. Major anxiousness, restlessness, or in severe pain.
1. Emotionally agitated (e.g., mania, post surgery of craniotomy).
2. Speaks only foreign language and no interpreter available.
3. Sensory deprivation/overloaded (e.g., isolated unit, intensive care unit).
4. On complicated therapies/interventions such as urgent operation of tracheostomy in the unit.
5. Severe pain requiring frequent analgesia.

9a. Compliant.
1. Accepting self as in need of health care and following the treatment regime.

9b. Intermittent compliant.
1. Accepting the exam or treatment rendered but denying being ill or intermittently non-cooperative.

9c. Non-compliant.
1. Denies being ill and resistant to the treatment needed (e.g., pulls out
I.V.)

10a. Minor health knowledge deficit.
1. Physical check-up admission patient.
2. Routine admission patient for routine treatment (e.g., chemotherapy).
3. Readmitted patient who has been admitted for more than 24 hours and no further health education scheduled on classification day.
4. Newly diagnosed patient who has been admitted for more than 48 hours, and no further health education scheduled.

10b. Not ready for health education.
1. Agitated or severely anxious.
2. Denying having an illness.
3. In critical condition (e.g., massive bleeding).

10c. Moderate health knowledge deficit.
1. Readmitted new patient in the first 24 hours of admission.
2. Newly diagnosed patient in the 24-48 hours of admission.

10d. Ready for education.
1. Emotionally stable or asks questions related to disease.

10e. Scheduled education.
1. General health education about the disease will be provided to the patient on the classification day.

10f. Major health knowledge deficit.
1. Newly diagnosed patient and admitted within first 2 days.

11a. Family or friends available.
1. Relatives or friends with the patient at the time of admission.
2. Relatives or friends visited within 24 hours of the classification time.

11b. Limited SS.
1. No relative/friend (RF) at bedside at the time of admission.
2. No RF visited within 24 hours of the classification time.

11c. Stressful family situation.
1. Recent (in 1/2 year) or impending divorce.
2. Family has disease or disabled family member(s) in recent 3 months.
3. Patient or family abuses drugs/alcohol.

12a. Has health insurance for hosp.
1. Hospitalization bills paid by other groups (e.g., Medicare, Medicaid, research funding, charity institutions).

12b. Has no financial problem.
1. No complaint of the difficulty on self-paying bills for hospitalization.

12c. Has financial problem.
1. Complaint of the difficulty on self-paying bills for hospitalization.

13a. Minor bedside data collection.
1. One or two simple lab sample collection(s) a day (e.g., fluid profile data, sputum).
2. Vital signs q8h or less.
3. Intake & output q8h or less.

13b. Moderate bedside data collection.
1. Moderate complicated and consecutive lab samples collecting (e.g., glucose tolerance test (GTT), blood sugar check qid).
2. Three or more simple lab sample collecting; vital signs, EKG or intake & output q4h.

13c. Complicated bedside data collection.
1. Three or more lab sample collecting including the complicated and consecutive lab samples collecting.
2. Vital sign, EKG or intake & output q2h or more.

18. Vital signs (VS) or intake and output (I&O).
1. The frequency of VS or I&O is started by counting backwards 8 hours before the classification time and forward 4 hours after the classification time.

19a. Bedside exam or treatment <1 hour.
1. Simple exam at bedside such as soap enema.

19a. Exam or treatment at other department.
1. Not at unit because of exam or treatment on other department and will be transferred back to the unit before 8 p.m.

19b. Bedside complicated exam or treatment ≥1 hour but <2 hours.
1. Complicated exam or treatment at bedside such as lumbar puncture, abdomina-centesis.

19c. Bedside complicated exam or treatment ≥2 hours.
1. Complex exam or treatment at bedside (e.g., enema for hepatic coma, perineal dialysis).

Date _____
Rater Initials ____
(Researcher use only)

Patient Classification Form

Patient Demographic Data:

Hosp #_____ Major Diagnosis_____
Age_____ Sex_____ Race_____

Maximum Numbers of Patients: _____
Nursing time required for this patient_____
(Researcher Use only)

Pt Classification Scores: Scoring 1, 2, or 3 to each item

Item of Self-Care Requisites or Deficits	Score	Item of Self-Care Requisites or Deficits	Score
1. Air intake		9. Promotion of normalcy	
2. Fluid intake		10. Health knowledge	
3. Food intake		11. Social support	
4. Bladder elimination		12. Financial security	
5. Bowel elimination		13. Lab data collection	
6. Hazard prevention		14. Medical regimen	
7. Activity		15. Surgical regimen	
8. Rest			

Optional Classification Scores: Scoring 1, 2, or 3 to any items applicable to patient

Item of Self-Care Requisites or Deficits	Score	Item of Self-Care Requisites or Deficits	Score
16. Tracheal or endotracheal suctioning		18. VS or I&O	
17. Multiple I.V. lines		19. Special examinations or treatment procedures	

APPENDIX B

FIRST CONTENT VALIDITY PROCEDURE

Content Validity Procedure
The Orem Patient Classification System (OPCS)

This patient classification instrument is designed to be used by registered nurses to classify clinical patients based on patient requisites/deficits of self-care. This instrument is focused on the adult medical-surgical patient population between 18 and 65 years of age in an acute care clinical setting. The objective of the instrument is to measure nurses' workload in patients' care for the determination of appropriate allocation of nurse staffing.

This instrument consists of 15 fixed items based on Orem's (1991) framework of self-care requisites. Each item is followed by a group of indicators (the content domains of the interests) that identify individual patient's self-care agency in meeting his/her self-care requisites, or the therapeutic nursing technologies from authorized sources by which patient's self-care requisites can be met. A list of definitions of the indicators is attached to explain meaning of the indicators. Based on nursing workloads for each indicator in meeting individual's self-care requisites, these indicators are further classified into three modes of nursing care complexity: minimum, moderate, intensive (see appendix A), and a score of 1, 2, and 3 is assigned to each mode accordingly to give weights for nursing workloads. Additionally, there are 4 optional items used to give extra weights to certain indicators.

As a content expert in Orem's model and/or nursing service administration, you are asked to evaluate each item of the instrument in terms of the following questions:

1. Is the item congruent with the conception of self-care requisites presented by Orem (1991)?
2. Is the item relevant to the objective of the instrument for measuring the nursing staff workloads for patient care?
3. Do indicators of the item specifically measure self-care requisites?

A 4-point rating scale is applied to each one of the above 3 questions for every item on the attached "ITEM CONTENT ASSESSMENT SHEET." Please enter a check mark (✔) into the blank box on each rating scale that best represents the validity of the item/indicator. At the end of this assessment sheet, several lines of space are also provided for your recommendations.

Thanks for the opinions of your expertise.

Item Content Assessment Sheet

Directions: Please note **each** item for: its congruence with the framework of self-care requisites in Orem's (1991) theory; its relevance as a means of determining nursing staffing; and the specificity of the indicators in representing the specific self-care requisite.

Item 1:

Air intake

Congruence to Theory

Not Congruent — Highly Congruent

Relevance for Staffing

Not Congruent — Highly Congruent

Specificity to the Requisite

Not Congruent — Highly Congruent

Item 2:

Fluid intake

Congruence to Theory

Not Congruent — Highly Congruent

Relevance for Staffing

Not Congruent — Highly Congruent

Specificity to the Requisite

Not Congruent — Highly Congruent

Item 3:

Food intake

Congruence to Theory

Not Congruent — Highly Congruent

Relevance for Staffing

Not Congruent — Highly Congruent

Specificity to the Requisite

Not Congruent — Highly Congruent

Item 4:

Bladder elimination

Congruence to Theory

Not Congruent — Highly Congruent

Relevance for Staffing

Not Congruent — Highly Congruent

Specificity to the Requisite

Not Congruent — Highly Congruent

Item 5:

Bowel elimination

Congruence to Theory

Not Congruent — Highly Congruent

Relevance for Staffing

Not Congruent — Highly Congruent

Specificity to the Requisite

Not Congruent — Highly Congruent

Item 6:

Hazard prevention

Congruence to Theory

Not Congruent — Highly Congruent

Relevance for Staffing

Not Congruent — Highly Congruent

Specificity to the Requisite

Not Congruent — Highly Congruent

Item 7:

Activity

Congruence to Theory

Not Congruent — Highly Congruent

Relevance for Staffing

Not Congruent — Highly Congruent

Specificity to the Requisite

Not Congruent — Highly Congruent

Item 8:

Rest

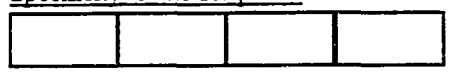
Congruence to Theory

Not Congruent — Highly Congruent

Relevance for Staffing

Not Congruent — Highly Congruent

Specificity to the Requisite

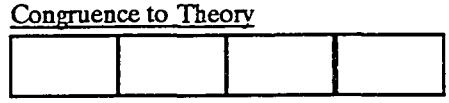


Not Congruent — Highly Congruent

Item 9:

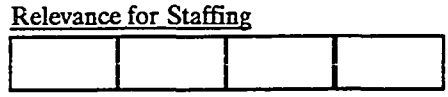
Promotion of normalcy

Congruence to Theory



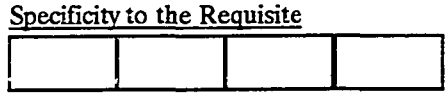
Not Congruent — Highly Congruent

Relevance for Staffing



Not Congruent — Highly Congruent

Specificity to the Requisite

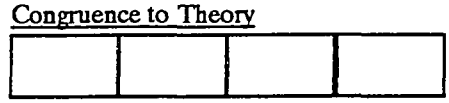


Not Congruent — Highly Congruent

Item 10:

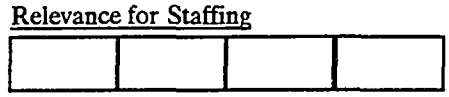
Knowledge about health state

Congruence to Theory



Not Congruent — Highly Congruent

Relevance for Staffing



Not Congruent — Highly Congruent

Specificity to the Requisite

Not Congruent — Highly Congruent

Item 11:

Social support (SS)

Congruence to Theory

Not Congruent — Highly Congruent

Relevance for Staffing

Not Congruent — Highly Congruent

Specificity to the Requisite

Not Congruent — Highly Congruent

Item 12:

Financial security

Congruence to Theory

Not Congruent — Highly Congruent

Relevance for Staffing

Not Congruent — Highly Congruent

Specificity to the Requisite

Not Congruent — Highly Congruent

Item 13:

Lab data collection

Congruence to Theory

Not Congruent — Highly Congruent

Relevance for Staffing

Not Congruent — Highly Congruent

Specificity to the Requisite

Not Congruent — Highly Congruent

Item 14:

Medical regimen

Congruence to Theory

Not Congruent — Highly Congruent

Relevance for Staffing

Not Congruent — Highly Congruent

Specificity to the Requisite

Not Congruent — Highly Congruent

Item 15:

Surgical regimen

Congruence to Theory

Not Congruent — Highly Congruent

Relevance for Staffing

Not Congruent — Highly Congruent

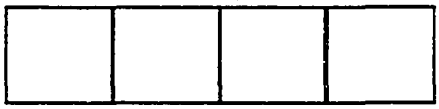
Specificity to the Requisite

Not Congruent — Highly Congruent

Item 16:

Frequent use of tracheal or endotracheal suctioning

Congruence to Theory



Not Congruent — Highly Congruent

Relevance for Staffing

Not Congruent — Highly Congruent

Specificity to the Requisite

Not Congruent — Highly Congruent

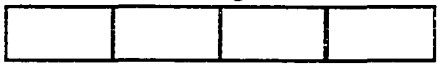
Item 17:

Multiple I.V. lines

Congruence to Theory

Not Congruent — Highly Congruent

Relevance for Staffing



Not Congruent — Highly Congruent

Specificity to the Requisite

Not Congruent — Highly Congruent

Item 18:

VS or I&O

Congruence to Theory

Not Congruent — Highly Congruent

Relevance for Staffing

Not Congruent — Highly Congruent

Specificity to the Requisite

Not Congruent — Highly Congruent

Item 19:

Special examinations or treatment Procedures

Congruence to Theory

Not Congruent — Highly Congruent

Relevance for Staffing

Not Congruent — Highly Congruent

Specificity to the Requisite

Not Congruent — Highly Congruent

Recommendations:

_____

_____

_____

_____

_____

APPENDIX C

INSTRUMENT AS REVISED FOR A SECOND ROUND EVALUATION OF CONTENT VALIDITY

Orem's Patient Classification System (OPCS)

The attached patient classification instrument consists of directions, guidelines for patient classification, a listing of medical-surgical patient classification indicators, definitions of the indicators, and a patient classification sheet. The instrument is designed for use by registered nurses to classify adult medical-surgical patients in an acute care clinical setting. The objective of the instrument is to measure nurses's workload needed in patients' care for the determination of appropriate allocation of nurse staffing.

Directions for Using the Orem's Patient Classification System (OPCS)

The OPCS was developed from a framework provided by Orem's Self-Care Model. The instrument is designed to assist Nurses in classifying the nursing care needs of adult medical-surgical patients in an acute care clinical setting. Once nursing care needs are classified the nursing workload required to provide patient care can be quantified and an appropriate allocation of nurse staffing can be determined.

The instrument consists of 18 items divided into three parts; Part A contains 11 items derived from universal self-care requisites, Part B contains three items related to health deviation self-care requisites, and Part C contains four optional items related to self-care deficits. Items on Part C identify individual patient self-care deficits that are selected from indicators within items on Part A and Part B. These items are included to provide extra weighing scores to the instrument when the indicators are applicable to an individual patient.

The indicators for each item were developed from 1) patient self-care agency for each universal self-care requisites, 2) patient self-care deficits identified from health-deviation self-care requisites of the individual patient when he/she becomes ill or injured, 3) health deviation self-care requisites arising from medical measures or treatment. All the indicators for each item were grouped into three modes of nursing care needs: minimum, moderate, and intensive. A score is assigned to each level as follows:

1 = Minimum nursing care assistance
2 = Moderate nursing care assistance
3 = Intensive nursing care assistance

Total scores possible range from 15 (minimal nursing care needed) to 54 (intensive nursing care needed). In situations where a patient has more than one indicator of nursing care needed, the higher score is to be assigned. For example if a patient does not show any sign of dehydration but does have a regular IV line, the score would be 2.

Definitions are provided for patient classification indicators of nursing care needed. Registered nurses should maintain familiarity with these definitions to insure consistent application of the tool. It is suggested that hospitals implement a patient classification audit system to maintain inter-rater reliability of the instrument above 90%.

Guidelines for Using the Orem's Patient Classification System (OPCS)

1. All hospitalized medical-surgical patients should be classified daily by their R.N. caregivers no later than 10:00 a.m.

2. Classification should include **all** patients:
   a. Present on the unit as of 10 a.m. and who have not been discharged.
   b. Who have been discharged but are still present on the unit.
   c. Off the unit for surgery, treatments, examinations or are on passes and will return by 6:00 p.m.

3. Do **not** classify patients:
   a. Discharged or transferred prior to 10:00 a.m.
   b. In surgery and will not return to the unit prior to 06:00 p.m.
   c. Off the unit for examinations, treatments or are on passes but will not return prior to 6:00 p.m.

4. Complete Demographic Data on the Patient Classification Form (PCF).

5. Obtain data from observation of the patient, the patient's chart and the shift-change nursing report.

6. Classify patients according to listing of the Orem's Patient Classification Indicators and the Definitions of Patient Classification Indicators.

7. Record score for each item on the PCF.

8. Enter the total patient classification score on the PCF.

9. Initial the PCF and then give it to the registered nurse on the unit who is responsible for data gathering.

Orem's Patient Classification Indicators

Part A: Categories of universal self-care requisites

Item	Nursing Care Needed		
	Minimal (Score=1)	Moderate (Score=2)	Intensive (Score=3)
1. Air intake	*No respiratory difficulty *Old trache-ostomy	*Some difficulty in respiration -Use of chest tube -Oral suction -Tracheal, or endotracheal (ET) suction x 1-2/shift	-Tracheal or ET suction x 3-4/shift or more often -Use of ventilator
2. Fluid intake	-No restriction on fluid intake	*Fluid restriction *Regular IV	*Critical IV
3. Food intake	-No nausea -No vomiting -No assistance with meals -NPO	-Nausea -Vomiting -Need assistance with meals -N-G tube drainage by gravity *N-G tube decompression *Minimum bloody vomit	*Tube or total feeding *Extensive bloody vomit
4. Bladder elimination	-Continent -Regular elimination	*Difficulty in urination *Use of catheter	*Urinary incontinent

Notes: * indicates definitions of terms provided on list.

5. Bowel elimination	-Regular bowel movements -Proficient self-care of colostomy -Continent	*Difficulty in bowel movements *Recent colostomy *Tube to abdomen *Incontinent *Minimum rectal bleeding	*New colostomy *Extensive rectal bleeding
6. Physical motion	-Ambulatory *Proficient use of prosthesis	*Weak motor strength *Local restraints *New use of prosthetics	-Absolute bed rest *Body restrained *Completely immobilized
7. Integrity of sensory organs	-No impairment or deficiency in sensory organs	*Some impairments in sensory organs	*Extensive deficits in sensory organs
8. Orientation	-Oriented	*Disoriented -Delusional	-Unconscious
9. Comfort	-Normal rest or sleep pattern -No discomfort identified	*Minor anxiety, depression, or restless *Moderate pain -Vertigo	*Major anxiety, depression, or restless *Severe pain
10. Social relations	-Self reliant *Cooperative - Able to interact with others	*Intermit-tently cooperative *low interaction with others	*Non-cooperative *Manipulative *Autistic *No or very low interaction with others
11. Health knowledge	*Adequate health knowledge	*Minor health knowledge deficit	*Major health knowledge deficit

Notes: * indicates definitions of terms provided on list.

Part B: Categories of health-deviation self-care requisites

Item	Nursing Care Needed		
	Minimal (Score: 1)	Moderate (Score: 2)	Intensive (Score: 3)
12. Information monitoring	*Minor bedside lab data monitoring -Vital signs (VS) less than q4h -Intake and output (I&O)/shift or less	*Transfer in *Moderate bedside lab data monitoring -VS q4h -I&O q4h	*Newly admitted *Complex bedside lab data monitoring -VS q2h or more -I&O q2h or more
13. Delivery of medicine	-Oral medications -Injection medications	*IV medications ≤2 doses/shift	-IV medications ≥3 doses/shift
14. Skin care	-Skin intact -Postoperative >48 hrs -No change dressing (CD) -CD qd	-Post operation >24 hrs -Tube to wound (other than chest & stomach) -CD bid	*Operative day -Postoperative ≤24 hrs -CD tid or more

Notes: * indicates definitions of terms provided on list.

Part C: Optional categories of self-care deficits:

Item	Nursing Care Needed		
	Minimal (Score: 1)	Moderate (Score: 2)	Intensive (Score: 3)
15. Frequent tracheal suction	-Tracheal or endo-tracheal suction x 5-6/shift	-Tracheal or endotracheal suction x 7-8/shift	-Tracheal or endotracheal suction x 9/shift or more often
16. Multiple IV line	-Two IV lines	-Three IV lines	-Four IV lines or more
17. * Vital signs (VS) or intake and output (I&O)	-VS qh -I&O qh	-VS q15' plus q30' ≥2 hours -I&O q30'	-VS q15' ≥2 hours
18. Special examinations or treatment procedures	*Exam or treatment at other dept *bedside simple exam or treatment	*Bedside urgent simple operation -Bedside simple exam or treat x 2 or more *Bedside compound exam or treatment	*Bedside urgent complex operation -Bedside compound exam or treatment x 2 or more *Bedside complicated exam or treatment

Notes: * indicates definitions of terms provided on list.

NOT FOR USE BY DATA COLLECTOR:

1) Total scores is the sum of all scores obtained from each item in the columns.

2) Minimum score is 14 and maximum score is 54.

Definitions of Patient Classification Indicators

1a. No respiratory difficulty.
1. Normal respiration.
2. Stoma in trachea.

1b. Old tracheostomy.
1. Intermittent use of cap on the site of tracheostomy.
2. Use of tracheostomy but no suction needed.

1c. Some difficulty in breathing.
1. Shortness of breath.
2. O_2 prn use.
3. Productive cough.
4. Frequent non-productive cough.

2a. Fluid restriction.
1. Limit or force fluids.

2b. Regular IV.
1. IV with plain fluids or drugs such as insulin, bicarbonate.
2. Heparin lock IV.

2c. Critical IV.
1. IV with blood transfusion or life-saving drugs such as anti-hypotension/hypertension or chemotherapy agents.
2. IV infusions requiring special catheter to heart.
3. Specially designed IV lines, such as arterial line, CVP line, ICP line.

3a. Need assistance with meals.
1. Needs assistance with cutting meat, pouring cream and sugar, and so on.

3b. Minimum bloody vomit.
1. Bloody vomit approximate to 10 ml/3 hours or less.

3c. Tube feeding.
1. Use of N-G tube, G tube, or TPN tube for feedings of nutrients.

3d. N-G tube decompression.
1. Use of N-G tube connected with decompression suction machine.

3e. Extensive bloody vomit.
1. Bloody vomit approximately more than 10 ml/3 hours.

4a. Difficulty in urination.
1. Complaint of difficulty, frequency, burning, and/or urgency in urination.

4b. Use of catheter.
1. Use of condom, intermittent, or indwelling catheter.

4c. Urinary incontinent.
1. Inability to control the voiding of urine and no urinary catheter used.

5a. Difficulty in bowel movements.
1. Abdominal distention, constipation, fecal impaction, diarrhea.

5b. Recent colostomy.
1. Colostomy of more than 3 days and not proficient in self-care.

5c. Tube to abdomen.
1. Tube to surgical wound, sore, or cavity to withdraw body fluids or discharges.

5d. Incontinence.
1. Involuntary passage of feces and flatus.

5e. Minimum rectal bleeding.
1. Rectal bleeding approximate to 10 ml/h or less.

5f. New colostomy.
1. Have colostomy <3 days.

5g. Extensive rectal bleeding.
1. Rectal bleeding approximately more than 10 ml/h.

6a. Proficient use of prosthesis or aids.
1. Proficient use of prosthesis or aids for motor activity (e.g., wheelchair, crutches; artificial limbs).

6b. Weak motor strength.
1. Side-rails up.
2. Bed rest with bathroom privilege.
3. Up with assistance (e.g., Post-OP days, complications of DM neuropathy or CHF).
4. Impaired range of motion (stiffness on parts of body) such as PP cast, cranio-fixation, hemiplegia, or paraplegia.

6c. Local restraints.
1. Use of extremity restraints.

6d. New use of prosthesis.
1. New prosthesis not used proficiently.

6e. Body restrained.
1. Use of vest restraint.

6f. Completely immobilized.
1. Cranio-traction.
2. Physically unresponsive or only responsive to pain and pressure (e.g., unconscious, quadriplegia).

7a. Some impairments in sensory organs.
1. Impaired hearing (use of hearing aid) or seeing (e.g., snowflake cataract, excluding use of contact lenses or eye glasses).

7b. Extensive deficits in sensory organs.
1. Deaf, blind, or almost blind (can see only lights and no figures), mute.

8a. Oriented.
1. Alert about time, person, or place.

8b. Disoriented.
1. Confused about time, person, or place.

9a. Minor anxiety, depression, or restless.
1. Expresses anxiety, depression, or restless but is not agitated (e.g., fear of lost job, fear of the coming special treatment).
2. Insomnia.
3. Speaks foreign language and some local language.
4. Speaks only foreign language and interpreter available.

9b. Moderate pain.
1. Complains of intermittent pain or discomfort; needs little or no pain medicine.

9c. Major anxiety, depression, or restless.
1. Expresses anxiety, depression, or restless and is emotionally agitated (e.g., mania, S/P craniotomy).
2. Speaks only foreign language and no interpreter available.
3. Sensory deprivation/overloaded (e.g., isolated unit, ICU).
4. On complicated therapies/interventions such as urgent operation of tracheostomy on the unit.

9d. Severe pain.
1. Severe pain requiring frequent analgesia.

10a. Cooperative.
1. Accepts need for health care and collaborate with others to his/her treatment regime.

10b. Intermittently cooperative.
1. Accepts exam or treatment rendered but denies being ill or intermittently non-collaborative.

10c. Low interaction with others.
1. Stays alone unless urged by others to interact and/or fail to initiate verbal interaction.

10d. Non-cooperative.
1. Denial of being ill and resistant to the treatment needed (e.g., self-removal of IV).

10e. Autistic.
1. Has self-centered trends of thought or behavior which are not corrected with external information.

10f. Manipulative.
1. Uses others for selfish purposes, disregarding their individuality, integrity, and rights.

10g. No or very low interaction with others.
1. Not able to interact with others.
2. Avoids or refuses most interactions or activities, urged by others.

11a. Adequate health knowledge.
1. Physical check-up admission patient.
2. Routine admission patient for routine treatment (e.g., chemotherapy).
3. Readmitted patient in 2nd week of hospitalization.
4. Newly diagnosed patient in 3rd week of hospitalization.

11b. Minor health knowledge deficit.
1. Readmitted new pt in 1st week of hospitalization.
2. Newly diagnosed patient in 2nd week of hospitalization.
3. Health education for patient and/or family is planned to give on the day of classification, contents to include 1-2 topics (e.g., self injection, self bladder catheterization, self selection of special diet).

11d. Major health knowledge deficit.
1. Newly diagnosed patient in the 1st week of hospitalization.
2. Health education for patient and/or family is planned to give on the day of classification, content to include 3 or more topics.

12a. Minor bedside lab data monitoring.
1. Collecting one or two simple lab samples for lab data monitoring qd or less (e.g., fluid profile data, sputum).

12b. Transfer in.
1. Transferred into the unit before 10 a.m. of the classification date.

12c. Moderate bedside lab data monitoring.
1. Collect consecutive lab samples for lab data monitoring (e.g., glucose tolerance test, blood sugar check qid).
2. Three or more simple lab samples collecting.

12d. Newly admitted.
1. Arrived on unit before 10 a.m. of the day.

12e. Complex bedside lab data collecting.
1. Collect simple lab sample collecting and consecutive lab samples for lab data monitoring.

13a. IV medications
1. Intravenous infusion of medicines to patient including setting up morphine pump which provides limited dosage of morphine q15' when pumped by patients.

14a. OP day.
1. In surgery after 8 a.m. and will return to unit before 6 p.m.

17a. Vital signs (VS) or intake and output (I&O).
1. The frequency of VS or I&O is counted backward 4 hours before the classification time and forward 4 hours after the classification time unless specified.

18a. Exam or treatment at other department.
1. Not on unit for exam or treatment on other department; transfer back to the ward before 8 p.m.

18b. Bedside simple exam or treatment.
1. Simple procedure for examination or treatment, such as plain or soapy water enema, Fleet enema, pre-op skin preparation, insertion of N-G tube, or Foley's catheter.

18c. Bedside simple urgent operation.
1. Simple urgent operation at bedside, such as insertion of arterial line, CVP line, or endotracheal tube.

18d. Bedside compound exam or treatment.
1. Moderate complex exam or treatment at bedside, such as abdomina-centesis, chest centesis, analysis of gastric juice, lumbar puncture.

18e. Bedside urgent complex operation.
1. Complex urgent operation at bedside, such as tracheostomy.

18f. Bedside complicated exam or treatment.

1. Complicated exam or treatment at bedside, such as cleansing enema for hepatic coma, peritoneal dialysis.

Patient Classification Form (PCF)

Please provide patient information in the following sequences:

1) enter the patient's demographic data in area provided.
2) estimate the maximum number of patients you could take care of per shift if all the patients were like the one you are classifying. Round one position after decimal point.
3) score the patient according to the directions and guidelines provided in the OPCS.

Demographic Data:

Hosp #_____ Major diagnosis_____
Age_____ Sex_____ Race_____
Maximum number of patients: _____
Nursing time required for this patient_____

(Researcher use only)

Scores: 1 = minimum nursing care needed
2 = moderate nursing care needed
3 = intensive nursing care needed

Score Item

Part A: Universal self-care requisites

_____ 1. Air intake
_____ 2. Fluid intake
_____ 3. Food intake
_____ 4. Bladder elimination
_____ 5. Bowel elimination
_____ 6. Physical motion
_____ 7. Integrity of sensory organs
_____ 8. Orientation
_____ 9. Comfort
_____10. Social relations
_____11. Health knowledge

Part B: Health deviation self-care requisites

_____12. Information monitoring
_____13. Delivery of medicine
_____14. Skin care

Part C: Optional items of self-are deficits

_____15. Tracheal suction
_____16. Multiple I.V. line
_____17. VS or I&O
_____18. Special examination or treatment procedures

Total Score: _____ **Rater Initials:** _____

APPENDIX D

SECOND CONTENT VALIDITY PROCEDURE

Secondary Content Validity Procedure
The Orem's Patient Classification System (OPCS)

The OPCS was developed from Orem's (1991) framework of 1) universal self-care requisites that are common to all human beings and 2) health-deviation requisites that may exist when the individual is ill/injured, or when measures for the diagnosis and/or treatment of health care deficits are used.

Several procedures were used to develop the OPCS. First, items of the instrument were developed by the focus of patient/nursing agency from the framework of universal self-care requisites and the framework of health-deviation requisites that arose from medical measures. Thus, Part A of the instrument contained 11 items for universal self-care requisites, which were arranged by the sequence of physical, psychological, and sociological needs of the individual patient. Part B of the instrument contained three items for health-deviation self-care requisites that arose from medical measures.

Secondly, each item developed examples of indicators which represent 1) patient self-care agency on each universal self-care requisite, 2) self-care deficits converted from universal self-care requisites when an individual patient has disabilities/defects because of being ill or injured, or 3) health deviation self-care requisites that arise from medical measures/treatments. All the indicators were further organized into three different modes of nursing care needed; minimum, moderate, and intensive, and a score of 1, 2, and 3 was assigned to each mode accordingly, based on the anticipated nursing workloads for each mode.

Thirdly, four optional items were selected from indicators of self-care deficits on Part A and B, and a group of indicators for the particular self-care deficit were developed and organized following the same procedure as were mentioned for Part A and B to provide extra weighing scores for the instrument when these extra indicators were applicable to the individual patient. As a result, this instrument consists of 18 items: 14 fixed items and 4 optional items (Please see attachment: The Development of Patient Classification Items from Orem's Model).

Finally, the directions and the guidelines for patient classification, and the patient classification sheet were developed. Attached is the instrument titled as The Orem's Patient Classification System (OPCS) for reference.

As an expert in Orem's model and/or nursing service administration, you are asked to evaluate each item of the instrument in terms of the following questions:

1. Is the item congruent with the concept of self-care requisites presented by Orem (1991)?
2. Is the item relevant to the objective of the instrument for measuring the nursing staff workloads for patient care?
3. Does the indicators of the item specifically measure a self-care requisite?

A 4-point rating scale is applied to each of the above 3 questions for every item on the attached "ITEM CONTENT ASSESSMENT SHEET." Please evaluate each item along these three dimensions by placing a check mark (✔) into the appropriate box. If you have further recommendations, please place them at the end of this assessment sheet.

Thank you for your expertise.

References

Orem, D. E. (1991). Nursing: concepts of practice (4th ed.). St. Louis: Mosby.

Item Content Assessment Sheet

Directions: Please note **each** item for: its congruence with the framework of self-care requisites in Orem's (1991) theory; its relevance as a means of determining nursing staffing; and the specificity of the indicators in representing the specific self-care requisite.

Item 1:

Air intake

Congruence to Theory

Not Congruent — Highly Congruent

Relevance for Staffing

Not Congruent — Highly Congruent

Specificity to the Requisite

Not Congruent — Highly Congruent

Item 2:

Fluid intake

Congruence to Theory

Not Congruent — Highly Congruent

Relevance for Staffing

Not Congruent — Highly Congruent

Specificity to the Requisite

Not Congruent — Highly Congruent

Item 3:

Food intake

Congruence to Theory

Not Congruent — Highly Congruent

Relevance for Staffing

Not Congruent — Highly Congruent

Specificity to the Requisite

Not Congruent — Highly Congruent

Item 4:

Bladder elimination

Congruence to Theory

Not Congruent — Highly Congruent

Relevance for Staffing

Not Congruent — Highly Congruent

Specificity to the Requisite

Not Congruent — Highly Congruent

Item 5:

Bowel elimination

Congruence to Theory

Not Congruent — Highly Congruent

Relevance for Staffing

Not Congruent — Highly Congruent

Specificity to the Requisite

Not Congruent — Highly Congruent

Item 6:

Physical motion

Congruence to Theory

Not Congruent — Highly Congruent

Relevance for Staffing

Not Congruent — Highly Congruent

Specificity to the Requisite

Not Congruent — Highly Congruent

Item 7:

Integrity of sensory organs

Congruence to Theory

Not Congruent — Highly Congruent

Relevance for Staffing

Not Congruent — Highly Congruent

Specificity to the Requisite

Not Congruent — Highly Congruent

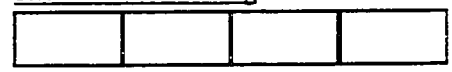
Item 8:

Orientation

Congruence to Theory

Not Congruent — Highly Congruent

Relevance for Staffing



Not Congruent — Highly Congruent

Specificity to the Requisite

Not Congruent — Highly Congruent

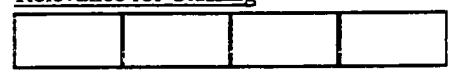
Item 9:

Comfort

Congruence to Theory

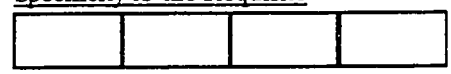
Not Congruent — Highly Congruent

Relevance for Staffing



Not Congruent — Highly Congruent

Specificity to the Requisite

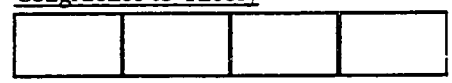


Not Congruent — Highly Congruent

Item 10:

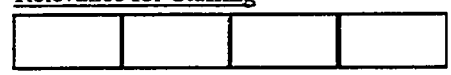
Social relations

Congruence to Theory



Not Congruent — Highly Congruent

Relevance for Staffing



Not Congruent — Highly Congruent

Specificity to the Requisite

Not Congruent — Highly Congruent

Item 11:

Health knowledge

Congruence to Theory

Not Congruent — Highly Congruent

Relevance for Staffing

Not Congruent — Highly Congruent

Specificity to the Requisite

Not Congruent — Highly Congruent

Item 12:

Information monitoring

Congruence to Theory

Not Congruent — Highly Congruent

Relevance for Staffing

Not Congruent — Highly Congruent

Specificity to the Requisite

Not Congruent — Highly Congruent

Item 13:

Delivery of medicine

Congruence to Theory

Not Congruent — Highly Congruent

Relevance for Staffing

Not Congruent — Highly Congruent

Specificity to the Requisite

Not Congruent — Highly Congruent

Item 14:

Skin care

Congruence to Theory

Not Congruent — Highly Congruent

Relevance for Staffing

Not Congruent — Highly Congruent

Specificity to the Requisite

Not Congruent — Highly Congruent

Item 15:

Tracheal suction

Congruence to Theory

Not Congruent — Highly Congruent

Relevance for Staffing

Not Congruent — Highly Congruent

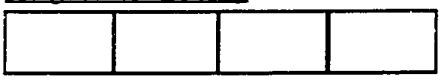
Specificity to the Requisite

Not Congruent — Highly Congruent

Item 16:

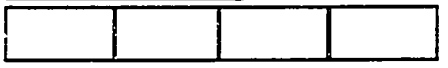
Multiple IV lines

Congruence to Theory



Not Congruent — Highly Congruent

Relevance for Staffing



Not Congruent — Highly Congruent

Specificity to the Requisite

Not Congruent — Highly Congruent

Item 17:

VS or I&O

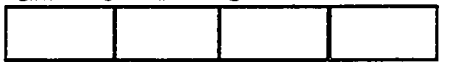
Congruence to Theory

Not Congruent — Highly Congruent

Relevance for Staffing

Not Congruent — Highly Congruent

Specificity to the Requisite



Not Congruent — Highly Congruent

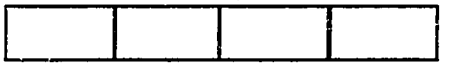
Item 18:

Special examinations or treatment

Congruence to Theory

Not Congruent — Highly Congruent

Relevance for Staffing



Not Congruent — Highly Congruent

Specificity to the Requisite

Not Congruent — Highly Congruent

Recommendations:

_____

_____

_____

_____

_____

_____

GRADUATE SCHOOL
UNIVERSITY OF ALABAMA AT BIRMINGHAM
DISSERTATION APPROVAL FORM

Name of Candidate Pi-Ru Chiu

Major Subject Nursing Service Administration

Title of Dissertation Development and Testing of a Patient Classification Instrument Derived from Orem's Model

Dissertation Committee:

_______________, Chairman

____

____

____

____

Director of Graduate Program _______________

Dean, UAB Graduate School _______________

Date 12/15/94

PS 1426