1984

An Assessment Of The Reliability And Validity Of Phaneuf Nursing Audit Scores.

Robert William Vogler
University of Alabama at Birmingham

INFORMATION TO USERS

This reproduction was made from a copy of a document sent to us for microfilming. While the most advanced technology has been used to photograph and reproduce this document, the quality of the reproduction is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help clarify markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure complete continuity.

2. When an image on the film is obliterated with a round black mark, it is an indication of either blurred copy because of movement during exposure, duplicate copy, or copyrighted materials that should not have been filmed. For blurred pages, a good image of the page can be found in the adjacent frame. If copyrighted materials were deleted, a target note will appear listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed, a definite method of "sectioning" the material has been followed. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For illustrations that cannot be satisfactorily reproduced by xerographic means, photographic prints can be purchased at additional cost and inserted into your xerographic copy. These prints are available upon request from the Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases the best available copy has been filmed.

University
Microfilms
International
300 N. Zeeb Road
Ann Arbor, MI 48106

Vogler, Robert William

AN ASSESSMENT OF THE RELIABILITY AND VALIDITY OF PHANEUF NURSING AUDIT SCORES

The University of Alabama in Birmingham D.S.N. 1984

PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark ✓.

1. Glossy photographs or pages _____
2. Colored illustrations, paper or print _____
3. Photographs with dark background _____
4. Illustrations are poor copy _____
5. Pages with black marks, not original copy _____
6. Print shows through as there is text on both sides of page _____
7. Indistinct, broken or small print on several pages ✓
8. Print exceeds margin requirements _____
9. Tightly bound copy with print lost in spine _____
10. Computer printout pages with indistinct print _____
11. Page(s) _____ lacking when material received, and not available from school or author.
12. Page(s) _____ seem to be missing in numbering only as text follows.
13. Two pages numbered _____. Text follows.
14. Curling and wrinkled pages _____
15. Other Dissertation contains pages with print at a slant, filmed as received.

AN ASSESSMENT OF THE RELIABILITY AND VALIDITY OF PHANEUF NURSING AUDIT SCORES

by

ROBERT WILLIAM VOGLER

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, University of Alabama at Birmingham

BIRMINGHAM, ALABAMA

1984

ABSTRACT OF DISSERTATION
GRADUATE SCHOOL, UNIVERSITY OF ALABAMA IN BIRMINGHAM

Degree D.S.N. Major Subject Nursing

Name of Candidate Robert William Vogler

Title An Assessment of the Reliability and Validity of Phaneuf Nursing Audit Scores

The purpose of the study was to generate estimates of reliability and validity for the Phaneuf Nursing Audit instrument. The tool was used to rate 300 medical records of adult terminally ill cancer patients who expired during their final admission to a southeastern hospital during 1974-1979. A 63-case subsample was used to generate unweighted subscale scores. Inter- and intra-rater reliability was maintained above the .80 level throughout the data collection phase.

Audit results were consistent with findings available in the literature. Nurses documented technical aspects of care, observed symptoms and reactions, and consistently carried out physician's orders. Nurses failed to record basic documentation, initiate and maintain assessments and care plans on patients. Emotional aspects of care involving patients and families were not well documented. Patient teaching, promotion of health and physical activities were poorly documented. Quality profile scores

for the sample resulted in a rating at the good level. The sample was found to be representative on the basis of age and sex using MANOVA (p=.05). Tentative support of moderately acceptable reliability and statistically significant validity was found. The Alpha coefficient for the total scale was .85. Estimates of validity utilizing the regression statistic produced significant F statistics for all seven subscales (p=.05).

Recommendations include: (a) replication of the study with other patient groups across settings and over time; (b) identification of baseline quality profiles and monitoring changes in institutions providing nursing care; (c) use of the instrument in nursing education programs to assess the practice of students, (d) for nursing service administration to monitor and enforce policies and procedures concerning nursing documentation; and (e) provision of continuing education for practicing nurses who care for terminally ill cancer patients.

Abstract Approved by: Committee Chairman Anne E. Belcher

Program Director

Date 11/30/84 Dean of Graduate School

ACKNOWLEDGEMENTS

This endeavor became a reality because of the support and guidance given the researcher by faculty, peers, family, and friends. I would like to extend my appreciation to Dr. Marie O'Koren, Dean of the School of Nursing and a valued committee member. Without her foresight, the doctoral program in nursing and this study would not have been possible. Her direct and incisive mind is appreciated. Also a thanks goes to Dr. Jean Kelley, Associate Dean for Graduate Programs in the School of Nursing. Her administrative gudiance and written communication have directed my navigation through the academic waters.

Dr. Anne Belcher, committee chairperson, has my appreciation and gratitude. I am thankful for her unfailing support and sense of humor that sustained me throughout this research endeavor. Dr. Belcher's work with oncology patients has proved invaluable. To Drs. Carl Miller and Phyllis Horns, a thanks for their careful review of the manuscript and consistently valuable suggestions.

A special note of appreciation goes to Dr. Joanne Scungio, my committee statistician, who was incisive and always constructive with her criticism. Dr. Scungio's research with pediatric oncology patients and their

families provided valuable insight into these research efforts. I hope that I may continue to be infused with her committment to the research process and be able to emulate her tenacity in face of life's adversity.

A thanks to Dr. Martha Hedley, whose eye for format and syntax is without peer. Her knowledge has substantially eliminated the redundancy of this process for me and provided direction.

I would like to express my appreciation to two of my peers with whom I journeyed through the academic labrinyth. To Dr. Mable Lamb, a loyal and valued friend, thanks, and to Dr. Lillian Ericksen, a commrade and valued person, also a thanks.

To my family and friends I offer my gratitude and appreciation for each of them being who they are in my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The passage of Public Law 92-603 in 1972, which required the monitoring of health care, has served as an impetus for the nursing profession to scrutinize more closely the health care services provided to the consumer. Chance (1980) identified several additional factors that have led to the current interest in nurses' evaluating and monitoring health care. These factors include: (a) the limited availability of financial resources, (b) effects of legislation and practice acts on the regulation of services, and (c) the increasing responsibility of the profession to supervise and improve services.

Nursing as an evolving health care profession has responded to these societal forces by setting practice standards, developing measurement tools, and conducting a limited amount of nursing evaluation research. The American Nurses' Association, as the major professional organization, developed practice standards by 1973. A small number of measurement or evaluation instruments have been constructed (Phaneuf, 1972 & 1976; Carter, Hilliard,

Castles, Stoll, & Cowan, 1976; Hegyvary & Haussman, 1975) in an attempt to determine the quality of nursing care. Initial efforts by individual nurse researchers investigating nursing care have focused on organizational structure (Jelinek & Dennis, 1976), nursing process (Wandelt & Stewart, 1975), and patient outcomes (DeGeyndt, 1960, Lindeman, 1972).

An early effort in nursing evaluation research by Phaneuf (1966) resulted in the development of a process-oriented audit instrument. The tool is used for the retrospective documentation of seven nurse functions in the patient's medical records. After the cycle of nursing care has been completed, the patient's chart is audited using the Phaneuf instrument. A score is generated for each of the seven functions. The extent to which nurses have used the Phaneuf Nursing Audit (Phaneuf, 1972 & 1976) is not clearly documented in the literature, but the developer indicates (Phaneuf, Note 1) that it is widely used in evaluating the care provided by nurses in community public health agencies and, to a lesser degree, in hospital-based care settings. Isler (1974) reports utilization of the audit tool in a university medical center. Until the publication of a study by Ventura (1980), no significant amount of information had been available in the literature concerning estimates of reliability and validity for scores generated by the Phaneuf Nursing Audit instrument.

Basic psychometric properties which directly relate to the credibility of research findings, such as reliability and validity, have been routinely neglected by nurse researchers. Goodwin & Prescott (1981) surveyed 61 nursing research instruments and found that 32 did not report reliability. Estimates of reliability and validity are generally understood to be germane elements of research efforts and must be calculated each time an instrument is used by the investigator. If estimates addressing the error variance and construct validity are not calculated then little confidence can be placed in the research findings. Utilization of such research findings in decision making can not be done with any degree of certainty. It was the need to document the reliability and validity of an existing instrument of measure that prompted this study.

Polit & Hungler (1978) clearly identify the need for estimates of reliability and validity as essential standards for assessing the quality of an instrument used in any research efforts. Kerlinger (1973) further states:

> There is growing understanding that all measuring instruments must be critically and empirically examined for their reliability and validity. The day of tolerance of inadequate measurement has ended. (p. 473)

Need for the Study

The results of a study carried out by the Western Interstate Comission for Higher Education in 1974 (Krueger,

Nelson, & Wolanin, 1978) used the Delphi technique to ascertain priorities for nursing research. The panel concluded:

> . . . that the determination of valid and reliable indicators of quality nursing care had been rated first in priority by the whole panel for impact on patient care by clinicians for value to the profession. On this basis, the measurement of quality of care was established as the priority for the targeted research effort. (p. 84)

The current literature supports this priority of Krueger and associates. "In essence nursing has a minimal number of instuments most of which need considerable validity and reliability testing. . . (Atwood, 1980, p. 104). Bloch (1975) urges the establishment of a clearinghouse for nursing practice evaluation that would provide a critical analysis for published and unpublished research efforts in relation to validity and reliability. It is imperative that nurse researchers utilize the same accepted psychometric standards common to other fields that produce scientifically based knowledge.

Purpose of the Study

The purpose of this study was to generate estimates of reliability and validity for the Phaneuf Nursing Audit instrument. A modified form of the Phaneuf Nursing Audit was used to make comparisons to the original instrument.

Research Questions

The following questions were posed:

1. Are the mean subscale scores of the sample significantly different on the basis of age and sex?

2. What is the internal consistency of all the items of the Phaneuf Nursing Audit instrument?

3. How predictive of the total scores are the subscale scores of the Phaneuf Nursing Audit instrument?

4. What comparisons can be made concerning dichotomous scores generated by the Phaneuf Nursing Audit and by continuous scores generated by a modified form of the Phaneuf Nursing Audit?

5. What constructs can be derived from a modified form of the Phaneuf Nursing Audit?

Definition of Terms

The following operational definitions were accepted by this investigator for use in this study:

1. Reliability - the degree of consistency or accuracy with which an instrument measures an attribute as measured by Coefficent Alpha.

2. Validity - the degree to which an instrument measures what it is constructed to measure and in this study measured by the regression statistic.

3. Audit - a retrospective evaluation of nursing care based on the documentation recorded in the patient's medical records as measured by the Phaneuf Nursing Audit.

Assumptions

The following assumptions were identified for this study:

1. Estimates of validity and reliability are legitimate concerns of nurse researchers.

2. Indicators of specific process components of nursing care can retrospectively be identified and measured from documentation found in patients' medical records.

3. Nurses document nursing care rendered to hospitalized patients during the care cycle.

Limitations

The following limitations were considered applicable for this research:

1. Data collected represent a retrospective audit and are dependent upon the amount of documentation recorded on the patient's medical records.

2. Appraisal of nursing care in this study is limited to the seven nurse functions utilized by Phaneuf in the nursing audit scale.

3. Due to the setting, sample size, and admission criteria, findings of this study are not representative of other patient populations or settings.

CHAPTER II

REVIEW OF THE LITERATURE

This chapter presents a review of the literature relevant to three areas of the study: psychometric theory concerning the concepts of reliability and validity; consideration of a framework into which this descriptive research could be placed, including the area of evaluation research and more specifically the process component identified by Donabedian (1966); and finally, a review of the literature relating specifically to the Phaneuf Nursing Audit. A description of the statistics used for analysis in this study is included in this chapter. The chapter concludes with a summary.

Psychometric Theory

Reliability and validity are two statistical concepts basic to the development of measurement instruments. Although both the concepts of reliability and validity are necessary in order for a measure to meet psychometric standards, Best (1977) indicates that the reliability of an indicator must be addressed before consideration is given to the validity of a tool.

The reliability of an instrument is judged by how consistently the tool produces a stable measure of the variable it was constructed to measure. Classic test theory indicates a score or a measure results from two ingredients: (a) a true score, or the amount of true score present, and (b) measurement error, or the amount of measurement error present. A reliable measure would be one that contains as small an amount of measurement error as possible while producing as much of a true score as possible each time the same subject is tested (Cozby, 1977; Nunnally, 1967).

The score generated by an instrument represents the sum of two parts: (a) a true score and (b) an error component. Scores are never totally error free, and since it is the error in measurement that affects reliability, the error component of scores is of vital interest. Error may result from numerous sources, such as the subjects, the raters, the instrument and/or the administration of the instrument (Waltz & Bausell, 1981).

Psychometric literature identifies three aspects of reliability that are usually considered: (a) stability, (b) internal consistency, and (c) equivalence. Fox (1970) specifies four techniques utilized for estimating reliability: (a) test-retest, (b) alternate form, (d) split-half or odd-even, and (d) use of Kuder-Richardson procedures (mathematical extensions of the split-half

method). For example, if the instrument being assessed for reliability consists of 50 single items, the use of each item as an instrument should be reliable and behave in the same way as the total 50-item instrument. The Kuder-Richardson estimate is based on the "relationship of the response pattern to each separate item to the data from performance on the total instrument" (Fox, 1970, p. 242).

Reliability may also be assessed by computing a correlation coefficient (r) indicating the strength of relationship between two variables. In instances in which measures produce scores, the strength of r would indicate the consistency with which the instrument produces like scores for individuals over a period of time, or, how consistently the measure acts. The reliability coefficient identifies the amount of true score variance, as opposed to error variance. To the degree that measurement error is absent, the reliability coefficient will approach its maximum theoretical value of +1.0. Conversely, an unreliable measure will have a coefficient approaching zero, the theoretically minimum value. Consequently, a low reliability coefficient indicates that a measure is producing unstable or inconsistent scores (King, 1979). Although authors differ on acceptable coefficents, Fox (1970) indicates that an r of .70 is generally accepted as a minimal level of reliability, whereas an r of .80 is the minimum standard for purposes of evaluation. He strongly

suggests that evaluation research should accept the more stringent standard when dealing with reliability estimates.

A reliability estimate applies to the one-time use of the instrument with a selected sample under conditions specified by the methodology of the study. However,

> Researchers frequently and incorrectly assume that a reliability estimate obtained in one study applies to all uses of the instrument, even when the subjects and circumstances are quite different from those that yielded the reliability coefficient (Goodwin & Prescott, 1981, p. 324).

Additional concerns that must be addressed in relation to the statistical estimation of reliability involve the linearity and homeoscidasity of the data generated as well as the adequacy or representiveness of the data of the population from which the sample was drawn. The simple "eyeballing" of the data visually presented is done in order to ascertain whether the hypothesis concerning linearity is present, an important first step. If linearity of the sampled data is questionable, further examination emphasizing greater precision is necessary (Minium, 1970).

Homeoscidasity, or the equal variability of data along the line of best fit, should be inspected, and then investigated if necessary. If one or both of the distributions used to compute the line of best fit are skewed, then the resulting distributions do not fit the

correctness of the underlying linear hypothesis, and should be questioned and subsequently treated with other appropriate statistical techniques (Nunnally, 1967).

Another consideration is that the data sampled should fit a bell-shaped normal distribution (Minium, 1970). The requirement of normal distribution fit is extremely critical if any inferences are to be made concerning the population value of the coefficient. Otherwise, the value of r will vary from sample to sample and therefore will produce variance depending on the chance factors dealing with sample selection.

A further limitation imposed upon a correlation drawn from a specific population is the size of the sample from which the correlation was obtained. In general, small sample studies tend to produce less stable values of r and may not represent the true population value. Large sample studies tend to produce values of r which are similar from sample to sample since large samples generally produce smaller variances. Thus an r value resulting from a large sample will probably be more representative of the true population value. In an effort to produce more stable results, Minium (1970) supports the generation of large sample studies when producing estimations of reliability.

Reports concerning estimates of reliability coefficients should include a detailed description of the instrument used and the methodology applied in order to produce or

obtain the statistics. Since research reports may or may not be applicable to other research situations, research results are considered tentative and subject to confirmation in the research environment presently under investigation. They must be considered in the context of the conditions under which the estimates were produced for each study (Carmines & Zeller, 1979).

Instruments that use raters to generate scores are open to variability. The reliability of the instrument may vary on the basis of how consistently the raters apply the tool. Two aspects of variance in score production which involve raters are intra-rater reliability and inter-rater reliability. These are primary concerns for any research endeavor in which two or more raters apply a single instrument in order to generate scores (Horn, 1980).

Intra-rater reliability addresses the consistency with which one rater assigns scores with the same instrument on two different occasions. In contrast inter-rater reliability is the consistency with which two raters assign scores using the same tool. Inter-rater consistency should be considered when two or more raters are responsible for determining scores with a common instrument. Pearson's correlation coefficient is suggested by Waltz and Bausell (1981) to estimate reliability for an instrument used by two or more raters to produce scores. Nunnally (1967), however, purports the use of coefficient alpha to assess

the consistency with which raters assign scores using a rating instrument. The interpretation is not that the raters assign the same scores but rather that raters assign a matching relative order of scores. A +1.0 would be interpreted as complete agreement in the relative order of assigning scores for raters.

A commonly accepted procedure to control for inter-rater consistency is for the raters to assign scores with the instrument and then, two weeks later, rescore the same individuals' records. A measure of agreement is then determined between the raters two separate scorings by computing a Pearson Product Moment Correlation coefficient (Waltz & Bausell, 1981). A zero would represent complete inconsistency whereas a +1.0 would indicate complete consistency. A coefficient of .80 or above is generally acceptable in evaluation research (Carmines & Zeller, 1979).

Inter-rater reliability estimates should be employed during the training of raters in order to ascertain whether raters are consistently using the same or common criteria for scoring items on the instrument. Throughout the period of data collection, raters should assess and review the same records. A coefficient of .80 or above should be maintained throughout the data collection phase. If inter-rater agreement falls below the .80 level, data collection should cease and retraining should be undertaken with the

instrument. A level of .80 or above must be reestablished in order that a high degree of consistency is maintained throughout the data collection phase (Nunnally, 1967).

Waltz & Bausell (1981) indicate that the method of assessing reliability depends on several factors, including the type of measure, the phenomenon being studied, cost, logistics, time, and, most importantly, the availability of or the amount of previous estimates of reliability for the instrument. It is necessary to apply a more stringent measure of reliability if previous assessments are not available to the researcher. Until estimates become readily available in the literature the more stringent standard should be utilized.

Procedures for assessing reliability fall into three basic categories: (a) test-retest, (b) parallel forms, and (c) internal consistency (Polit & Hungler, 1978). Test-retest considers the consistency with which a measure generates scores for the same group of subjects on two different occasions. Parallel forms require the use of alternate forms of the same instrument given to one group of subjects on one testing occasion.

The procedure for determining internal consistency is more involved than that for test-retest and parallel forms. The consistency of performance on each individual item of the instrument is statistically compared to the total score. The tool is used to obtain scores with a

representative group sample on one occasion during the same administration and a statistical estimate of reliability is computed. A coefficient alpha value is computed for the individual scores. The alpha coefficient measures the extent to which performance on any one item is a good indicator of performance on any other item of the same instrument (Brown, 1976). It is generally the preferred measure of internal consistency because it generates a single value for a set of test data. Alpha is equal to the mean of the distributions of all possible split-half correlation coefficients associated with a particular group of data (Nunnally, 1978, Carmines & Zeller, 1979).

Several factors are important when using coefficient alpha. The alpha statistic is affected by the length of the intstrument and stability may be increased substantially with an increase in instrument length. Alpha obtained in one situation should not be applied to other situations. The size of alpha is directly related to the instrument variance. The larger the test variance, the larger the alpha value. The alpha statistic is not appropriate for a criterion referenced framework. It is the preferred estimator of reliability in norm referenced instruments. Norm referenced instruments measure a specific characteristic while identifying subjects possessing differing amounts of the characteristic (Waltz & Bausell, 1981).

After the measure of an instrument's reliability has been demonstrated, the next task is to determine validity. Batey (1970) notes that, "While reliability has been dealt with quite directly in much of the research literature, considerable hedging and just plain neglect can be found for the concept of validity" (p. 7).

Nunnally, (1967) identifies three types of validity: (a) face/content, (b) predictive, and (c) construct. Statistical measures are not required to assess face and/or content validity, since they are considered a lower level estimate of validity. Face validity refers to the appearance of the instrument and only requires superficial inspection in order to determine if the items are common to the domain being sampled. For example, if an instrument purports to measure nursing aility, the instrument would be expected to include items concerning nursing. Content validity is more expansive and usually requires that a panel of experts (or qualified idividuals in the domain of the content area of interest) review the items on the test and then judge whether each of the items represents the appropriate domain. Content validity is important for instruments designed to measure cognition and is especially significant during the phase in which the instrument development takes place. The usual practice is to develop many items related to the topic of the study, and then use random sampling to select a subset of cases for use in

constructing an instrument. At least two specialists should review the objectives in conjunction with the selected items in order to ascertain whether the selected items accurately represent the behavior of interest.

Estimates of higher levels of validity require differing statistical treatments. Criterion-related validity is referred to as either predictive or concurrent. Predictive validity estimates are obtained by administering an instrument to a group and later comparing their success on some independant criterion. Concurrent validity indicates the measure's ability to discriminate between individual scores and the subject's success on a selected criterion of interest at a specific point in time (Nunnally, 1967).

Other procedures used in assessing validity are contrasted groups, the experimental manipulation approach, and the multitrait-multimethod approach. The contrasted group approach is accomplished by identifying two groups, one known to be high-performers and the other theoretically low-performers in a specific domain of interest. The experimental manipulation technique involves the use of the theory or rationale underlying the construction of a measure to predict how a group of individuals will perform under induced conditions. Individuals are placed in controlled conditions and then administered the instrument. If individuals behave in a manner consistent with prediction, one can offer this as evidence for the presence

of contruct validity. If inconsistent behavior results and the prediction is not upheld, then the same steps are taken as was for the known groups method (Waltz & Bausell, 1981). The items on the tool must undergo revision and then be subjected to the test of prediction.

The multitrait-multimethod approach is based on two premises. First, convergent validity assumes that different measures of the same construct should correlate highly with each other, and secondly, discriminant validity infers that measures of different constructs should have a low correlation with each other. These approaches deal with two types of variance. The first type deals with the variability in a set of scores resulting from individual differences in subject's ability to respond appropriately to the trait being measured, or trait variance. The second type of variance relates to variability resulting from the particular type of measure used, or method variance. The size of a correlation between two measures is a function of both trait and method variances (Nunnally, 1967).

Construct validity, although developed several decades ago (Cronbach & Meehl; 1955, Campbell, 1966), has become more readily accessible to the individual researcher since the use of the computer. Construct validity addresses whether the instrument actually does measure the variable or construct it purports to measure. This type of validity is evaluated by determining the degree to which explantory

concepts or constructs account for performance on a measure or instrument (Carmines & Zeller, 1979). Studies of construct validity require that the researcher state a hypothesis suggesting the behavior of persons scoring either high or low on the measure. The researcher would then gather data to test this hypothesis. If the data do not support the hypothesis, additional testing is necessary.

From the statistical results of the data, the researcher is able to infer the validity of constructs supposed to be measured by the tool. If the data can be explained on the basis of the rationale underlying the constructs of the instrument, validity is said to be supported. However, should the data fail to support validity, revision of the instrument is necessary and new evidence is required to establish validity for the revised instrument (Nunnally, 1967).

The regression statistic is frequently used to analyze relationships between dependent variables and a group of independent variables. Generally the assumption is that the variables consist of either interval or ratio data but the regression statistic is robust enough to treat other types of data. The purpose of the statistic may be either to describe or infer the strength of relationship(s) based on the line of best fit projected through a group of data. Used as a descriptive tool it can be ascertained whether

there is linear dependence of one variable on other variable(s). Descriptively the regression statistic can be used to predict from variables, examine relationships between variables or examine causal theory (e.g., the magnitude or direction of the direct and indirect influences of variables). Inferentially regression may be used to assess relationships in a targeted population on the basis of sample data. For the purpose of generalizing from sample to universe the sample size should be greater than 200 in order for the statistic to approximate a normal distribution (Kerlinger & Pedhazur, 1973). Regression may be used inferentially to estimate and test hypotheses. Sample data are used to infer whether the observed linear association is statistically significant. The statistic produced is the F ratio that is then compared to tabled values. Confidence levels may be established by use of the Beta statistic. The regression statistic was used in this study to assess the relationship between total scores and subscales scores produced by the Phaneuf Nursing Audit.

Multiple analysis of variance (MANOVA) is a statistical extension of analysis of variance (ANOVA). ANOVA is a basic inferential statistic used to generalize from a sample group to a population. ANOVA in its simplest form is a test that determines if significant differences exist between the means of groups or scores. The raw data are used to calculate a value. The calculated value from the

data is compared at a pre selected level of significance (p=.05 or p=.01) to tabled values. The results allow determination of whether there is a significant difference between the mean scores of the groups or groups of scores. The accurateness of infering from a sample to a population assumes that the sample is randomly selected and is representative of the population under investigation.

If there is a significant difference, it can be concluded that the sample scores are not representative of the population and any subsequent statistical inference would be unjustified. Further statistical treatments are required if differences are found between groups or mean scores. Additional computations are available for within and between group influences. Use of ANOVA is not theoretically sound for making inferences when there is more than one dependent variable. MANOVA is the suggested statistical technique used to deal with two or more dependent variables. The statistic is adjusted for the means of variables that might not equal zero and the bias that may exist with more than one dependent variable (Winer, 1971, Huck, Cormier, & Bounds, 1974).

Hoetelling's T^2 procedure is commonly used to determine if there is a significant difference between the means of the dependent variables. If the calculated T value does not exceed the critical value at a predetermined level of significance then it can be inferred that differences do

not exist on the basis of the variables and the sample represents a multivariant population. Further statistical treatments can then be applied to the scores of the sample and inferences made to the population on the basis that the sample is representative of the population. T values are usually transformed into F values that are easily compared to available tabled values. MANOVA was used in this study to determine if the sample scores were representative on the basis of age and sex.

The statistical techniques of factor analysis are frequently used to: (a) reduce data, (b) identify underlying concepts represented by the manipulated data, (c) confirm or test the number of factors present, and (d) identify the statistical loading of factors. Therefore, factor analysis is a viable method to determine the validity of items on an instrument. Its use in this study was to confirm or test the number of categories present in the modified Phaneuf tool used to produce unweighted scores for the subsample. Rao's Canonical factoring, as discussed in the SPSS X: Users Guide (Statisticl Package for the Social Sciences X, 1983), supports the classical-factor model and assumes that the correlation matrix represents a sample of the cases from a population.

Conceptual Framework

A commonly accepted frame of reference for evaluation research has been developed by Donabedian (1966). He

identifies three aspects for conceptually evaluating health care: (a) structure, (b) process, and (c) outcome. These three dimensions are viewed as interrelated components in evaluating the care patients receive in the health care system. The three aspects of evaluation research are represented in Figure 1.

Structure is concerned with the setting in which the care is given. Studies focusing on this aspect may analyze organizational structure, facilities, equipment and/or personnel. Outcomes are usually criteria established before care is given and evaluated at some end point after the care has been rendered to patients.

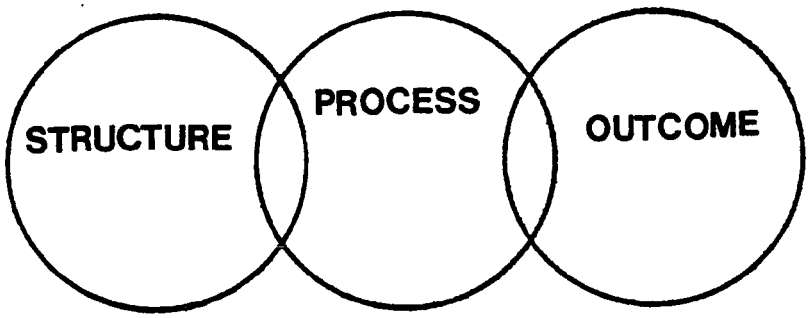

Figure 1. Donabedian's framework for evaluating health care.

Process Research

The aspect of particular interest for this research was the process of care, which can be either concurrent or retrospective. In concurrent studies data are collected while the patient is still actively receiving care, whereas

retrospective efforts focus on collecting data after the care has been provided and the patient has been discharged from the hospital. A study may focus primarily on one aspect, such as the treatment, or actual activities that were performed or carried out for the patient. Information on this process is usually collected by direct observation or with a form of chart or record audit (Staropoli & Waltz, 1970).

Donabedian (1980) suggests that the process element is advantageous while assessing the quality of care. Practitioners are able to identify certain criteria which then serve as expected standards of care. Documentation is usually readily available in patient's medical records and can be easily audited. Standards can be revised as treatment and care modalities change on the basis of any newly acquired knowledge. Process evaluation represents a readily available, accessible amount of information that can be monitored while ascertaining the level of care being provided to patients (Donabedian, 1980).

The major limitation in the use of process evaluation for the assessment of quality of care is the lack of scientific basis for much of accepted practice. The use of prevalent norms of professional practice must continually be updated to avoid perpetuation of the status quo. Process evaluation has focused mainly on documentation as the data base (Donabedian, 1982).

Inherent in the assessment of hospital records are certain limitations. A sample drawn on the basis of primary discharge diagnosis or operative procedure presents the following problems: (a) a segment of possible sample may be missing due to not receiving a specific diagnosis and therefore being excluded, (b) a sample drawn on the basis of diagnosis may not represent a probability sample because it is so homogeneous and may not represent the total possible population, and (c) the validity of the data retrieved from the medical records is contingent upon the accuracy and completeness of the material in the record. Donabedian (1976) indicates that a low level of correlation has been found for physicans between hospital practice and recording in medical records.

An additional concern is that the criteria used for selection of the sample do not account for the presence of other diagnosis(es) in addition to the primary diagnosis. A method of control for this factor is dependent upon the size of the sample drawn and whether it is large enough to control for the influence of additional diagnosis(es). Rendundancy could also be a problem. For example, if a task is done more than once, is that better? - and how many times is considered to be superior or harmful?

The assumption made regarding patients in critically ill conditions is that professional personnel would be performing the majority of tasks for the patient population

and subsequently documenting the activities of nursing care. This assumption is not, however, supported or evidenced in the literature (Donabedian, 1980).

In process evaluation criteria or standards are used as a basis for making judgments about the level or quality of care. Attempts at evaluation research, although thought to be of more recent development, can in actuality be traced to nursing's modern founder Florence Nightingale. Grier (1981) reviews seven sudies completed by Nightingale during the years 1858-1867. These included such topics as: (a) evaluations of patient outcomes, (b) nursing interventions; (c) health of the British Army; (d) cost effectiveness of various hospial unit sizes, (e) the risk to nurses in hospitals, (f) quality of individual nurses, and (g) a comparative ealuation of care provided by different hospitals. Nightingale constructed and utilized a tool to rate monthly the nurse in personal character and "acquirements" during her period of service. A scale of excellent, good, moderate, and imperfect, as well as a zero category, was used to rate the nurses' activities.

Nightingale viewed the use of the chart audit in the research process as an appropriate endeavor for nurses to pursue. Froebe & Bain (1976) reinforce this idea by stating: "For nurses to measure nursing using a formal audit procedure is as appropriate as for physicians to measure medical practice" (p. 74).

More recently the Joint Commission on Accreditation of Hosptials (1970) has defined audit as a means of looking at the charts of discharged patients to determine the type of care they received, a retrospective review of the patient care process. The purpose of chart review is usually to identify deficiencies in the patients' condition or status and the aspects of nursing service that may have aided these deficiencies.

Bloch (1980), speaking to process evaluation, indicates that two tasks need to be especially addressed by nurse researchers. These are development of a measureable process criteria and development of reliable and valid methods, "for measuring the process of nursing care in all its various forms, including both the physical aspects and the process, as well as the psychosocial and cognitive aspects" (p. 258).

Galton and Reilly (1977) report a process research study using a chart audit to collect data concerning terminally ill cancer patients. The investigators ascertained whether these patients, during their final two weeks of hospitalization, received the same level of care as other acutely ill patients.

The checklist that was developed assessed the frequency in which the nurse recorded in the following categories; (a) pain, (b) comfort/mobility, (c) personalized reporting, (d) assessment, (e) protection, (f) skin/mouth, and (g)

nutrition/elimination. Ideal frequencies were determined by having 16 nurse experts validate the frequencies. The non-randomized sample of 20 patients was selected for inclusion in the study by identifying the most recently expired patients in each of two hospital settings.

The findings of Galton and Reilly's study included the following: (a) several categories exceeded the ideal frequencies; (b) protection and prophylactic care (including activities of skin/mouth care) scored 109.71% of the projected ideal frequency; and (c) nutrition and elimination exceeded the ideal frequencies. The lowest scores were in the category of patient comfort. These resulted in only a third of the recordings considered optimal. Personalized reporting resulted in 45% of the expected ideal frequency. The items condition and pain received 69.57% and 74.95% respectively of the ideal scores (Galton & Reilly, 1977).

The authors suggest that ideal frequencies may have been set too high or too low, depending on the category. Nurses may have neglected to record properly in those categories receiving low scores, or the category may not have been a need for these terminally ill patients. Categories receiving high scores may represent unnecessary redundance in nurse recordings or the setting of low ideal frequencies by the nurse panel of experts (Galton & Reilly, 1977).

Phaneuf Nursing Audit

When Phaneuf attempted to develop a nursing audit tool in 1952 there were no formal nursing methods for measuring the quality of care. The Phaneuf instrument was developed during the estalishment of a home care program in a New York visiting nurse project funded by Blue Cross. Phaneuf (1964) distributed an early version of the audit to "nursing leaders, physicians and social workers. There was agreement as to the potential for such an instrument, and many valuable suggestions were offered" (p. 42). The distribution of the audit to authorities in the field was a possible attempt to establish face and/or content validity.

In an article appearing in 1968, Phaneuf states "because the audit was not a research operation, it has not been possible to carry reliability and validity checks to the point we would have wished to reach and report" (p. 80). The author continues by explaining that the reason for not estimating validity and reliability was the lack of uniform standardized report forms. Donabedian, in the preface to the first edition of Phaneuf's book (1972) states, "needless to say, not many of her [Phaneuf] recommendations and views in this area are supported by hard research. They do, however, represent almost 20 years of carefully considered experience with the nursing audit" (p. xvi).

In an audit cited by Phaneuf (1969), 30 cases of persons with cerebral vascular incident and 30 cases of persons

with heart disease were reported. Findings for the patients with cerebral vascular disease revealed that six received good care, eight received incomplete care (care was good as far as it went), 11 received poor care, and four received unsafe care. The overall findings for the patients with heart disease were that none received excellent care, three received good care, five received incomplete care, and 16 poor care. Six received care that was classified as unsafe. The study indicated that the best executed function was the carrying out of physician's orders, the one dependent nursing function. Observation of symptoms and reactions was fragmentary. Recommendations for the participating agency included the following:

1. Use of the International Classification of Diseases in obtaining and recording primary and secondary diagnoses.

2. Increased emphasis on obtaining orders regarding all medications used by the patient, whether or not they were administered by the nurse.

3. Exploration of reasons why nurse-to-physician communications are incomplete and why nursing judgments are so seldom conveyed. The committee had expected to find physician-to-nurse communications less than adequate; an unexpected finding was that nurse-to-physician communications were an even greater problem.

4. Increased attention to assessment of vital signs, with recording and use of findings.

5. Establishment of a policy of requiring nursing assessment of the patient's physical and emotional condition at the time of admission and at the time of discharge.

6. Development of standards for charting, using the nursing audit schedule as a process guide (Phaneuf, 1969).

Jelinek and Dennis (1976), after an extensive review of the literature concerning evaluation research of the quality of nursing, found the heaviest emphasis on process measures. The authors conclude:

> As a group, except for the Phaneuf audit and the Medicus-Rush instrument, the process measurement devices lack rigorous statistical manipulation. Theoretical bases are discontinuous or nonexistent, but follow the nursing care model when they are found. The basis toward task and away from judgments or decisions of nurses is also evident. Validity and reliability issues have been infrequently addressed. The exception to these limitations is the Medicus-Rush and to some extent, the Phaneuf audit. (p. 196)

A 1978 study by Ventura and Hageman entitled "Testing of the Reliability, Validity, and Sensitivity of Quality of Nursing Care Measures: Final Report" failed to include any data estimating the validity and reliability for the Phaneuf Nursing Audit, one of the instruments used in the study. Ventura (1980) reports correlations between the Phaneuf Audit and the Quality Patient Care Scale. Intra-class correlations were used in estimating inter-rater

reliability and provide an index of agreement among the raters generating audit scores. For the Phaneuf audit seven of the eight categories met criteria to support evidence of inter-rater reliability. The categories Application/Execution of Physician's Legal Orders and Observations of Symptoms and Supervision of the Patient resulted in the highest correlation (.93), while Reporting and Recording was the lowest correlation (.73) (n=15).

Hsung, cited in Phaneuf (1976), reports on a five-year project utilizing the Phaneuf Nursing Audit tool at the University of Michigan Hospital in Ann Arbor. An audit committee was formed late in 1971 to identify the level of nursing care given in their hospital setting. This newly formed committee utilized the audit tool to rate two charts, and further attempted to revise some of the subcomponents on the Phaneuf Nursing Audit. After the total committee rated the same chart with a revised form, "the members learned a hard lesson; deletion or revision destroys the validity of the tool. Therefore the tool was accepted as published to be used in [their] setting" (p. 127).

Since members of the committee represented different clinical backgrounds, the orientation of members was of great importance to success in auditing. An overview of the work, with an explanation of the purpose and methodology, was given each new member. During

orientation, a new recruit generally did one audit for practice and the results were reviewed with either the chairman or the member who was being replaced on the committee. Audit results were compared with the actual audit done by a regular member. The steps were repeated, and gradually the recruit began to review two or three charts officially. New members also met with the chairman in mini workshops to audit a chart together. Charts selected represented hospital stays of 10 to 14 days, plus others of particular interest to the committee. The next step for the new members was to increase the number of charts audited until they were able to audit at least seven charts.

Approximately every three months the committee reviewed a copy of the same chart together. Components were discussed, written definitions used, and consensus reached. Inter-rater reliability was assumed to be affirmed by this procedure. To strengthen validity, each person audited the same chart prior to the meeting where results were compared and discussed. No statistical estimates of validity were reported. Data from audits were regularly discussed with special emphasis given to the consistent application of the tool. Recommendations were then made on the basis of the findings. A five percent random sample of monthly discharges was audited. Reports of committee audits included the following:

1. a statistical summary and quality profile;
2. a summary of basic identifying data;
3. remarks written by the committee;
4. recommendations from the audit committee;
5. summary of recommendations to date. (Phaneuf, 1976, p.132)

Computer generation of a data base facilitated the use of the audit tool and the reporting. By 1976, 1,800 records and 19 audit reports had been completed. Although the commmittee theorized that the length of stay for patients might influence care, this was not supported by audit results (Phaneuf, 1976).

Summary

Estimates of validity and reliability are accepted standards for researchers. Currently, there are few estimates of reliability and validity for nursing evaluation instruments. Although the Phaneuf Nursing Audit has been used to evaluate nursing care since its construction in the early 1960's, few estimates of validity and reliability are available in the literature for this instrument. The process component of evaluation research, as proposed by Donabedian, is utilized as a theoretical basis for this research. Evaluation research, thought to be a recent development, can be traced back to the founder of nursing, Nightingale, who as early as the 1860's addressed the evaluaion of nursing care.

CHAPTER III

METHODOLOGY

Discussed in this chapter is the methodology used for this study. Included is a description of the sample, setting, instrumentation, and consideration given to the protection of human subjects. Also included are the procedures utilized for the collection of data and analysis of data.

The Sample and the Setting

The sample for this study consisted of 300 randomly selected medical records of adult terminally ill cancer patients who expired during the years 1974-1979. The patients were hospitalized in an acute care setting of a university medical center located in the southeastern region of the United States. An additional subsample of 63 medical records was randomly drawn from the sample.

The university medical center is located in the inner city metropolitan area with a population of 899,000 residents and serves as a regional cancer treatment and referral center for several southeastern states. This center is actively engaged in health care research and

includes a research librarian who provides access to the medical records archives and monitors use of the records.

The initial identification of a patient population from which to draw an appropriate sample was obtained from the university medical center tumor board registry records. The tumor registry maintains a computer data base on each cancer patient cared for in the medical center. The following information was available on patients diagnosed as having cancer: (a) identifying information (demographic); (b) basic epidemological information; (c) date of diagnosis; (d) primary site; (e) histological type; (f) extent of disease or stage; (g) definitive treatment; (h) follow-up information (including tumor persistance, reoccurences and time lags, and quality of survival); (i) end results; (j) addresses; and (k) telephone numbers (for assistance in follow-up).

A computer list of 3,534 diagnosed cancer registry patients seen in the university medical center during the years 1974-1979 was obtained. The criteria for admission to the sample were applied to the prospective population list. In order to be considered for possible inclusion in the study the patient must have:

1. been 18 years of age or older at the time of final admission to the medical center;

2. been hospitalized in the medical center complex as an inpatient; and

3. subsequently expired as an inpatient during the final hospialization in the regional medical center.

Medical records were excluded from the study if:

1. the patients expired from causes other than their malignant disease process (i.e., auto accident, farm tractor accident, or other extraneous trauma); or

2. the medical record was grossly incomplete (e.g., lack of nurses' notes, diagnoses, physician orders, and/or end results).

Rating the Sample

A random table of numbers was used to select medical records for possible inclusion in the sample. If the record met the admisssion criteria, the record was given a case number and was randomly assigned to be audited by one of the two raters. Each twelfth chart entering the study was audited by both raters separately in order to maintain inter-rater reliability at the level of .80 or above.

Additionally, intra-rater reliability was maintained by having each rater audit the same record after a minimal two-day time lag. This process resulted in upholding the .80 or above correlation coefficient throughout the entire data collection process.

Instrumentation

The Phaneuf Nursing Audit (PNA) tool (see appendix A) was utilized to obtain the data for this study. The audit extended over a period of four months and 17 days,

commencing in December of 1980 and concluding in May of 1981. The tool is available in The Nursing Audit: Self Regulation in Nursing Practice (Phaneuf, 1976), and permission for use of the instrument was given (see Appendix B) by the publisher. In addition, a modified form of the Phaneuf Nursing Audit (MPNA) was used to audit a 63-case subsample drawn randomly from the 300-case test bed. The author was contacted and discussion occurred concerning modification of the scale for purposes of this study (Phaneuf, Note 1).

The PNA was initially developed by Maria C. Phaneuf in the early 1960's to measure the process component of evaluation, by retrospectively identifying nursing process in patient records after completion of the cycle of care. The audit utilizes the seven legal functions of nurses identified by Lesnik and Anderson (1955) and serves as a basis in scoring some 50 components of the nursing process. The intrument itself is based on subsets of the seven nursing functions commonly found in state statutes governing the licensure and practice of nurses. The seven common nursing functions, one dependent and six independent, are:

1. Application and execution of physician's legal orders (dependent function),
2. Observation of signs, symptoms, and reactions,
3. Supervision of the patient,

4. Supervision of those participating in care,

5. Reporting and recording,

6. Application and execution of nursing procedures and techniques, and

7. Promotion of physical and emotional health by direction and teaching.

The PNA provides a nominal scale for rating each of the seven functions, and results in a total overall numerical rating of the nursing care. The instrument is intended to measure the extent to which the seven nursing functions were carried out by the nurse for the patient. Criterion measures for each of the seven subscales are used to indicate the extent to which the function was performed (see Appendix C).

The subscale scores which result from the subscore components are identified in Table 1. The number of items, point value, and percent of points are listed for each subscale. In addition, a "does not apply item value guide" is used to adjust the final score (see Appendix D). The total score is multiplied by the weighted value appearing in the table for that number of does not apply items resulting from the rating. The adjusted scores represent dichotomous data.

The total numerical values are grouped into ranges that denote judgments of excellent, good, incomplete (good as far as it goes), poor, and unsafe (Phaneuf, 1969). The

scores for each of the ranges are interpreted in the following manner: excellent (161-200), good (121-160), incomplete (81-120), poor (41-80), and unsafe (0-40).

Table 1

Number of Items, Point Value, and Percent of Points for the Subscales of the Phaneuf Nursing Audit

Subscale	# of Items	Point Value	% of Points
I. Application and execution of physician's legal orders	6	42	21
II. Observation of symptoms and reactions	6	40	20
III. Supervision of the patient	7	28	14
IV. Supervision of those participating in care	4	20	10
V. Reporting and recording	5	20	10
VI. Application and execution of nursing procedures and techniques	16	32	16
VII. Promotion of physical and emotional health by direction and teaching	6	18	9
Totals	50	200	100

The Modified Phaneuf Nursing Audit (MNPA) contains the same seven subscales and the 50 items included in the

orginal scale. The distinguishing feature of the MPNA is the absence of the weighted score. Each of the 50 items has a range of 0-4 and a not applicable column. Zero equals no documentation and is considered unacceptable. Some or little evidence of documentation equals a score of one but is unacceptable. The score for more than some or little documentation, but less than the minimal expected documentation is two and is not acceptable. Three is the minimal expected documentation and is acceptable. Four is the score for full documentation and is acceptable.

The decision point on the 0-4 scale is at the level of three. Is the documentation less than minimal or is the documentation the minimal expected and acceptable? If the documentation is less than minimally acceptable then the next decision is, is there no evidence (0), some or little evidence (1), or more than some or little documentation, but less than minimally expected (2). At the decision point if it is decided that the record is minimally acceptable but not fully documented the item receives a score of three. The record if fully documented would receive a score of four.

The scores from the MPNA represent continuous data instead of the dichotomous type scores derived from the original weighted scales. Appendix E includes a representation of the MPNA scoring scale and the amount of score allocated at each of the points on the scale.

Protection of Human Rights

The rights of all subjects were protected in accordance with the regulation utilized by the University of Alabama at Birmingham. The research proposal was submitted to the Institutional Review Board for Human Use. Since these research efforts utilized medical records of deceased patients, the Institutional Review Board determined no risk to subjects as long as individual patients were not identified during the data collection process (See Appendix F). Permission to use these records was obtained from the medical records department of the medical center from which the data were collected. A case/code number was assigned to each patient's medical record that was admitted to the sample. No individual other than the investigator and the rater had access to the identification numbers. Only the code numbers were used through each phase of this research.

Data Collection

Data collection for this study involved the training of the raters, establishing inter-rater and intra-rater reliability, and the actual rating of the sample and subsample. The researcher and a graduate prepared nurse, who had previous experience with the Phaneuf Audit, scored the 363 medical records necessary for the study.

Rater training sessions were conducted between the rater and the researcher to familiarize the investigator

with the utilization of the instrument. Agreement on the objective cues for each of the subscales and individual items was derived through training in a service setting where the instrument was utilized to audit patient records. Each rater audited the same records. Ratings were compared and included the subscale scores and the total scores. Agreement or non-agreement was reached by referring to the medical record involved in order to document objectively the degree of evidence present for the cue or subscale.

Reliability Estimates

The next step in the data collection phase consisted of estimating inter-rater reliability by auditing medical records and obtaining alpha coefficients of .80 or above for the seven subscales of the tool. A subsequent session of training and agreement was necessary to attain this standard.

The same process was applied to the MPNA. Common definitions concerning the degree of documentation present were necessary with the MPNA before establishing acceptable inter-rater reliability.

Data Analysis

The data were analyzed using the following procedures:

1. The data from coding sheets were directly entered into a computer account via a keyboard. A printout of the data was obtained and verified against the orginal coding sheets. Statistical analysis was performed by computer

using the SPSSX: Users Guide (Statistical Package for the Social Sciences X, 1983). Unless otherwise indicated all statistical analysis applies to both the sample and subsample of this study.

2. Frequencies were calculated for all chart audit data.

3. Multiple analysis of variance (MANOVA) was used to determine if the sample scores were representative on the basis of age and sex.

4. The regression statistic was used to assess the predictive validity of the subscale scores and the total scores.

5. Cronbach's coefficient alpha value was used to generate estimates of reliability for each subscale score and total score.

6. Factor analysis was used to treat the scores of the subsample in order to determine item loading.

CHAPTER IV

FINDINGS

This chapter presents the findings of this research study. First, the sample is descibed; next, subscale scores are presented for the sample; a quality score distribution for the sample follows; and a summary of the sample data concludes this section. Next, the data from the subsample are presented, including description of the mean subsample scores. The research questions are then addressed and a summary concludes the chapter.

Three hundred medical records were audited using the Phaneuf Nursing Audit instrument for a group of terminally ill cancer patients during their final hospitalization. Estimates of reliability and validity were generated for the audit scores. An additional subsample was randomly drawn from the sample for the purpose of producing unweighted scores. These 63 records were audited utilizing a modified form of the Phaneuf Nursing Audit tool.

Description of the Sample

Each rater audited approximately one-half of the sample. The sample consisted of 35% blacks, 64.7%

Caucasians, and .3%, or 1 individual, classified as Oriental. Just under half (45%), or 135 cases, were female with the remaining 165 (55%) being male. Age of the sample ranged from 18-99 years. Four patients were in their ninth decade. One hundred and sixteen (116) patients were within the age of 60 to 70 years and comprised 38.6% of the cases. The length of terminal hospitalization ranged from 1-100 days. However, 27 or 9% of the subjects expired within the first day of hospitalization. An additional 20% (60) expired within three days of admission, and 40% (120) of the patients died within the first week. The mean length of stay was 16.5 days.

The cancer site varied from as many as 88 cases diagnosed with leukemia and lymphomas to as few as three cases with skin cancer. The largest group (93) were categorized as other and included such sites as brain, bone, and germinal reproductive tissue. Data in Table 2 depict the frequency and percent of cancer sites.

While 77 (26%) records were appropriately signed and dated by nursing personnel, 233 records (74%) were not appropriately signed and dated. Nursing entries that indicted the level of caregiver were found in approximately one out of every three records audited, or 105 of the cases. Nursing care plans were also included in only one-third (101) of the records. Although the patient population consisted of terminally ill patients, the nurse

during admission, failed to record the assessment of the physical condition in slightly over half (52.7%) of the

Table 2

Frequency and Percent of Cancer by Site

Cancer Site	f	%
Skin	3	1.0
Oral	12	4.0
Lung	35	11.7
Urinary	12	4.0
Prostate	12	4.0
Pancreas	17	5.7
Colon and rectal	22	7.3
Leukemia and lymphomas	88	29.3
Other (brain, bone & germinal reproductive tissue)	93	31.0
Missing (primary vs. secondary)	6	2.0
Totals	300	100.0

cases. Assessment of the emotional condition of the patient was documented even less frequently. In only 45 of the 300 cases, or 15% of the time, did the nurse record an emotional assessment. Data in Table 3 summarizes categories of these three nursing responsibilities by frequency and percent.

Table 3

Frequency and Percent of Three Documented Nursing Responsibilities

Nursing Responsibility	f	%
Entries Signed & Dated		
yes	77	25.7
no	223	74.3
Entries Indicate Level of Caregiver		
yes	105	35.0
no	195	65.0
Care Plan Recorded in Chart		
yes	101	33.7
no	199	66.3

Description of the Subscale Scores

The mean scores for each of the seven subscales are described first. The subscales are: (a) application and execution of the physician's legal orders, (b) observation of symptoms and reactions, (c) supervision of the patient, (d) supervision of those participating in care, (e) reporting and recording, (f) application and execution of nursing procedures and techniques, and (g) promotion of physical and emotional health by direction and teaching.

Subscale I: Application and execution of physician's legal orders. The mean scores for Subscale I indicated that physician orders were always complete as well as current, and the diagnosis for the patient sample was almost always complete. Data in Table 4 depicts the mean scores for items in Subscale I. Evidence that the nurse understood the cause and effect produced a mean score of 5.35 out of a total possible score of 7.0. The mean score for orders promptly executed was 5.03. Evidence that the

Table 4

Mean Scores for Items in Subscale I:
Application and Execution of Physician's Legal Orders

Item #	Item	Mean Score
1.	Medical diagnosis complete	6.98
2.	Orders complete	7.00
3.	Orders current	6.98
4.	Orders promptly executed	5.03
5.	Evidence nurse understood cause and effect	5.35
6.	Evidence nurse took health history into account	3.30

Note: Total possible score for all items in Subscale I=7.

nurse took the health history into account generated a mean score of 3.3, and was the lowest in this subscale.

Subscale II: Observations of symptoms an reactions. Data in Table 5 show the mean scores of items in Subscale II. Item 11, patient to his condition, was lowest for this subscale producing a score of 3.84, only .84 of a point

Table 5

Mean Scores for Items in Subscale II: Observations of Symptoms and Reactions

Item #	Item	Mean Score
7.	Related to course of above disease(s) in general	6.25[a]
8.	Related to course of above disease(s) in patient	6.09[a]
9.	Related to complications due to therapy	6.61[a]
10.	Vital Signs	6.06[a]
11.	Patient to his condition	3.84[b]
12.	Patient to his course of disease(s)	4.26[b]

[a] Total possible score =7 [b] Total possible score =5

above the uncertain category. The score for item 12, patient to his course of disease(s), was 4.26.

Subscale III: Supervision of the Patient. The mean scores for items in Subscale III are presented in Table 6. Three items in this seven-item subscale produced a mean

score of less than 1.0 (incomplete). Three additional items generated a mean score above 3.0. Item 19, interaction with family and with others considered, had the lowest mean score (.73) followed by item 13, evidence initial nursing diagnosis was made, (.84), and item 16,

Table 6

Mean Scores for Items in Subscale III: Supervision of the Patient

Item #	Item	Mean Score
13.	Evidence initial nursing diagnosis was made	.84
14.	Safety of patient	3.61
15.	Security of patient	3.60
16.	Adaptation (support of patient in reaction to condition & care)	.89
17.	Continuing assessment of patient's condition and capacity	3.31
18.	Nursing plans changed in accordance with assessment	1.83
19.	Interaction with family and with others considered	.73

Note: Total possible score for all items in subscale III=4.

adaptation (support of patient in reaction to condition and care) with a mean score of .89. These three items all scored below the 1.0 (uncertain) level. Safety of the

patient (item 14) had the highest mean score (3.61), followed by security of the patient, (item 15), with a mean score of 3.60 and continuing assessment of the patient's condition and capacity (item 17) with a mean score of 3.31.

Subscale IV: Supervision of those Participating in Care (except the physician). Data in Table 7 show the mean scores for Subscale IV. The mean scores for the four items

Table 7

Mean Scores for Items in Subscale IV: Supervision of those Participating in Care

Item #	Item	Mean Score
20.	Care taught to patient, family or others, nursing personnel	.17
21.	Physical, emotional, mental capacity to learn considered	.13
22.	Continuity of supervision to those taught	.13
23.	Support of those giving care	.11

Note: Total possible score for all items in Subscale IV=5.

in Subscale IV ranged from .11 to .17. Support of those giving care (item 23) scored .11 and (item 20) care taught to patient, family or others, nursing personnel resulted in a score of .13.

Subscale V: Reporting and Recording. Data in Table 8 depict the mean scores for items in Subscale V. The highest mean score in this subscale (3.6) was for item 25, essential facts reported to the physician. The lowest mean score (.09) was produced by item 27, patient or family

Table 8

Mean Scores for Items in Subscale V: Reporting and Recording

Item #	Item	Mean Score
24.	Facts on which further Care depended were recorded	2.4
25.	Essential facts reported to physician	3.6
26.	Reporting of facts included evaluation thereof	3.1
27.	Patient or family alerted as to what to report to physician	.09
28.	Record permitted continuity of intramural and extramural care	2.11

Note: Total possible score for all items in Subscale V=4.

alerted as to what to report to physician. The remaining three scores (items 24, 26, 28) ranged from 2.1 to 3.1.

Subscale VI: Application and Execution of Nursing Procedures and Techniques. Subscale VI contains 16 items and has more items than any of the seven other subscales.

Ten of the 16 items generated mean scores above 1.0. The mean item scores for this subscale ranged from .05 to 1.97. Items 39 and 41 produced mean scores of .05. Item 29, administration and/or supervision of medications, generated a 1.97 out of a possible 2.0 mean score. Items 38 and 44 produced scores of .10 and .11 respectively. The mean scores for Subscale VI are listed in Table 9.

Table 9

Mean Scores for Items in Subscale VI: Application and Execution of Nursing Procedures and Techniques

Item #	Item	Mean Score
29.	Administration and/or supervision of medications	1.97
30.	Personal Care (bathing, oral hygiene, skin, nail care, shampoo)	1.41
31.	Nutrition (including special diets)	1.51
32.	Fluid balance plus electrolytes	1.72
33.	Elimination	1.62
34.	Rest and sleep	1.46
35.	Physical activity	.81
36.	Irrigations (including enemas)	1.04
37.	Dressings and bandages	.96
38.	Formal exercise program	.10
39.	Rehabilitation (other than formal exercise)	.05
40.	Prevention of complications and infections	1.50
41.	Recreation, diversion	.05
42.	Clinical procedures - urinalysis, B/P	1.79
43.	Special treatments (e.g., care of tracheotomy, use of oxygen, colostomy or catheter care)	1.59
44.	Procedures and techniques taught patient	.11

Note: Total possible score for all items on Subscale VI=2.

Subscale VII: Promotion of Physical and Emotional Health by Direction and Teaching. All six items produced mean scores of less than 1.0, or the uncertain level. The lowest item, in this subscale, teaching promotion and maintenance of health (item 48), produced a mean score of .05, while emotional support to the patient (item 46) generated a mean score of .49. Item 49 (evaluation of need for additional resources) scored .46. Mean scores for the items in Subscale VII are presented in Table 10.

Table 10

Mean Scores for Items in Subscale VII: Promotion of Physical and Emotional Health by Direction and Teaching

Item #	Item	Mean Score
45.	Plans for medical emergency evident	.35
46.	Emotional support to patient	.49
47.	Emotional support to family	.28
48.	Teaching promotion and maintenance of health	.05
49.	Evaluation of need for additional resources (e.g., spiritual, social service, homemaker service, physical or occupational therapy)	.46
50.	Action taken in regard to needs identified	.33

Note: Total possible score for all items in Subscale VI=3.

Quality Score Distribution

The overall score of the 300-case audit resulted in a mean quality score of 121, which is at the lower limits of the good category, or one point above the incomplete category. The frequency and percent of records demonstrating quality scores are presented in Table 11. No quality scores were audited in the unacceptable category, whereas seven cases were assigned the excellent label. Sixty percent (181 cases) were classified as good, an additional 30% (91) were rated in the incomplete quality score category. Twenty-one records, or 7% of the sample, were labeled as poor.

Table 11

Frequency and Percent of Quality Score Distribution for the Sample

Quality Score	Range	n	% of Sample
Excellent	161-200	7	2.3
Good	121-160	181	60.5
Incomplete	81-120	91	30.2
Poor	41-80	21	7.0
Unacceptable	0-40	0	0.0
Totals			100.0

Comparison of the audit for the 300 cases to the profile of excellence (Phaneuf, 1976) is represented in Figure 3. The mean score for Subscale I, documentation of application and execution of physician's legal orders, was 35 and represents the top of the good category. Subscale II, documentation of observation of symptoms and reactions, the highest subscale mean score, was 33, and produced the only excellent mean score for any of the seven subscales.

Subscale III, supervision of the patient, was documented in the incomplete range with a mean score of 15. Subscale IV, supervision of those participating in care (except for the physician), was next to the lowest mean subscale score (.5), which is at the bottom of the unsafe category. This subscale included documentation in the patient record of teaching, supervision, and support for those participating in care.

Subscale V, reporting and recording, had a mean score of 11, which is at the top of the incomplete range. This Subscale dealt with the documentation by staff of the condition of the patient and the patient's disease.

Subscale VI, documentation of application and execution of nursing procedures and techniques, with a mean score of 18 fell in the incomplete range. Subscale VII included documentation related to promotion of physical and emotional health by direction and teaching, and was the lowest scoring category. A score of two was in the bottom

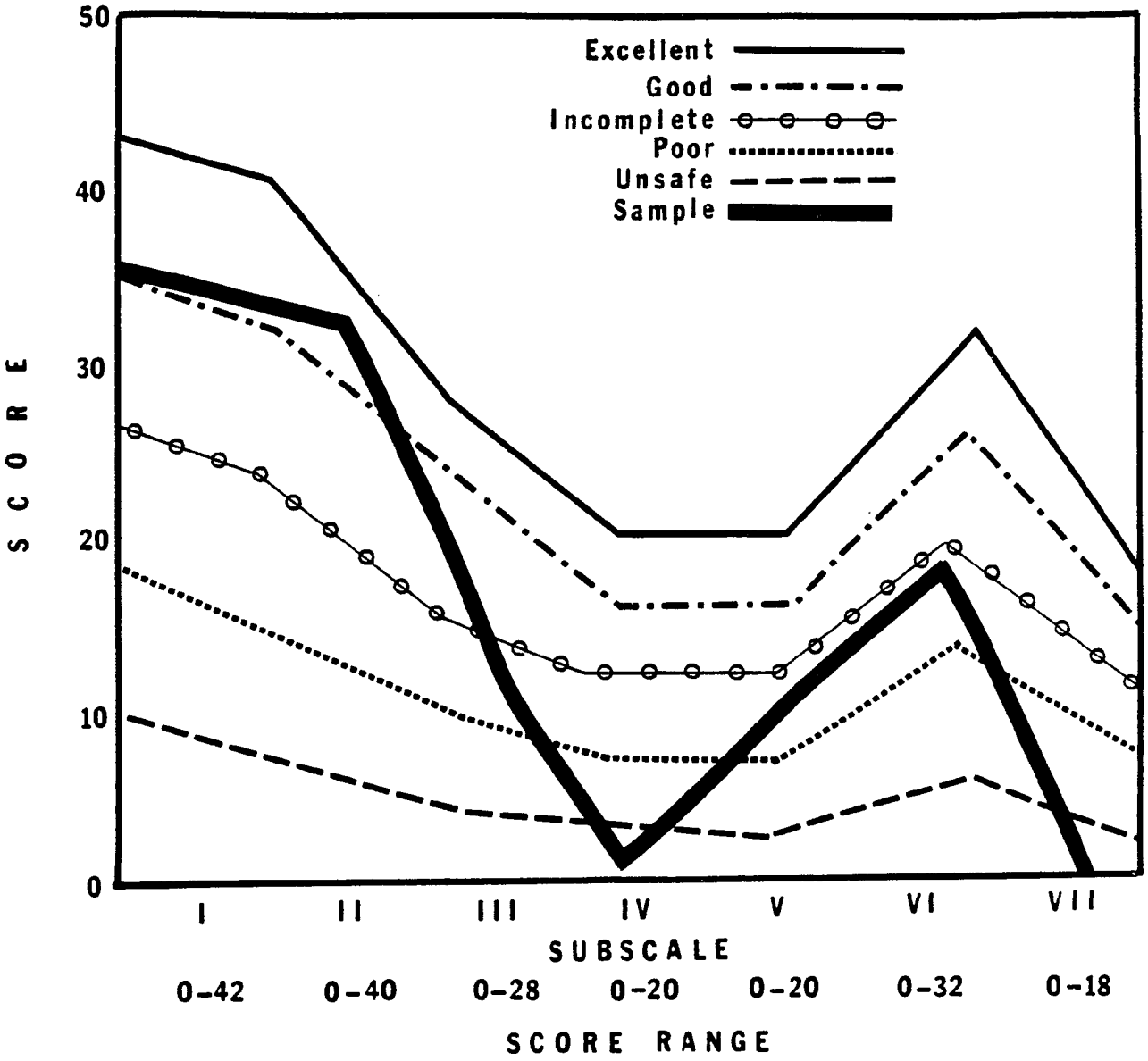

FIGURE 3. QUALITY SCORE PROFILE OF SAMPLE n=300

of the unsafe range. This unacceptable scoring category includes emotional support to the patient and family which indicates nurses infrequently documented this aspect of care in the patient records.

Description of the Subsample

Sixty-three records were randomly selected from the 300-case subsample and were rated using a modified form of the Phaneuf Nursing Audit (MPNA). Rater one scored 33 patient records and rater two rated 30 cases of the subsample. Unlike the scores from the original Phaneuf instrument, the resulting scores from the MPNA scale represented unweighted scores. It should be noted that the total possible score for each item on each subscale was consistently four.

Subscale I: Application and Execution of the Physician's Legal Orders. Mean scores for items in Subscale I are depicted in Table 12. The range of the mean scores for this six item subscale is from 1.61 to 3.98. The first three items of the subscale (medical diagnosis complete, orders complete, orders current) produced almost perfect scores of 3.96, 3.98, and 3.98, respectively (4.0 possible). Orders promptly executed (item 4) scored 2.90 and item 5 (evidence nursing understood cause and effect) scored 2.53. Item 6, evidence nurse took health history into account, was the lowest score for this subscale with a score of 1.61.

Table 12

Subsample Mean Scores for Items in Subscale I: Application and Execution of Physician's Legal Orders

Item #	Item	Mean Score
1.	Medical diagnosis complete	3.96
2.	Orders complete	3.98
3.	Orders current	3.98
4.	Orders promptly executed	2.90
5.	Evidence nursing understood cause and effect	2.53
6.	Evidence nursing took health history into account	1.61

Subscale II: Observations of Symptoms and Reactions.

Data in Table 13 show the mean scores of items in Subscale II. The range for the mean scores of this subscale is from 2.50 to 3.17. Only item 9 (related to complications due to therapy) had a mean score above 3.0 and was also the highest mean score for this subscale (3.17). Item 7, related to course of disease(s) in general, produced a 2.96 score. Items 10 and 8 (vital signs and related to course of disease(s) in patient) generated scores or 2.90 and 2.88, respectively. The lowest mean scores in this subscale were 2.50 for item 11 (patient to his condition), out of a possible score of 4.0, and 2.66 for item 12, patient to his course of disease(s).

Table 13

Subsample Mean Scores for Items in Subscale II: Observations of Symptoms and Reactions

Item #	Item	Mean Score
7.	Related to course of above disease(s) in general	2.96
8.	Related to course of above disease(s) in patient	2.88
9.	Related to complications due to therapy (each medication and each procedure)	3.17
10.	Vital signs	2.90
11.	Patient to his condition	2.50
12.	Patient to his course of disease(s)	2.66

Subscale III: Supervision of the patient. The mean scores of items in Subscale III are presented in Table 14. Mean scores in this subscale range from .50 for item 19 (interaction with family and with others considered) to 3.25 for item 15 (security of patient). Only two items in this 7-item scale produced a mean score of less than 1.0 (some or little evidence). Item 14 (safety of patient) and item 17 (continuing assessment of patient's condition and capacity) scored 2.93 and 2.95, respectively. Item 15, security of the patient, had the highest score (3.25) of any item in this subscale.

Table 14

Subsample Mean Scores for Items in Subscale III: Supervision of the Patient

Item #	Item	Mean Score
13.	Evidence initial nursing diagnosis was made	.79
14.	Safety of patient	2.93
15.	Security of patient	3.25
16.	Adaptation (support of patient in reaction to condition & care)	1.44
17.	Continuing assessment of patient's condition and capacity	2.95
18.	Nursing plans changed in accordance with assessment	1.60
19.	Interaction with family and with others considered	.50

Subscale IV: Supervision of Those Participating in Care (except the physician). Data in Table 15 depict the mean scores for Subscale IV. The range of the mean scores for this subscale is from .28 for item 20 (care taught to patient, family, or others, nursing personnel) to 3.52 for item 21 (physical, emotional, mental capacity to learn considered). The lowest mean score (.28) was produced by item 20, care taught to patient, family or others, nursing personnel, and was only slightly above the zero level of no documentation evident.

Table 15

Subsample Mean Scores for Items in Subscale IV: Supervision of Those Participating in Care

Item #	Item	Mean Score
20.	Care taught to patient, family, or others, nursing personnel	.28
21.	Physical, emotional, mental capacity to learn considered	3.52
22.	Continuity of supervision to those taught	3.42
23.	Support of those giving care	3.42

Subscale V: Reporting and Recording. Data in Table 16 depict the mean scores for Subscale V. Mean scores for the five items in this subscale range from .31 for item 27 (patient or family alerted as to what to report to physician) to 3.19 for item 25 (essential facts reported to physician). The highest mean score of 3.19 was for item 25, essential facts reported to the physician. The lowest mean score (.31) was produced by item 27, patient or family alerted as to what to report to physician. The remaining mean scores on items 24 (facts on which further care depended were recorded), 26 (reporting of facts included evaluation thereof), and 28 (record permitted continuity of intra-and extra-mural care) were 2.53, 2.90, and 2.53, respectively.

Table 16

Subsample Mean Scores for Items in Subscale V: Reporting and Recording

Item #	Item	Mean Score
24.	Facts on which further care depended were recorded	2.53
25.	Essential facts reported to physician	3.19
26.	Reporting of facts included evaluation thereof	2.90
27.	Patient or famiy alerted as to what to report to physician	.31
28.	Record permitted continuity of intramural and extramural care	2.53

Subscale VI: Application and Execution of Nursing Procedures and Techniques. The range of the mean scores of Subscale VI is .50 to 3.80. Item 29, administration and/or supervision of medications, generated the highest mean score (3.80) for this 16-item subscale. The next highest mean score was 3.20 (clinical procedures). Item 41, recreation/diversion, produced the lowest mean score of 0.50. Nine additional items (30, 31, 32, 33, 34, 36, 40, 43, 44) produced mean scores of 2.0 or above. The mean scores of items in Subscale VI are depicted in Table 17.

Table 17

Subsample Mean Scores for Items in Subscale VI: Application and Execution of Nursing Procedures and Techniques

Item #	Item	Mean Score
29.	Administration and/or supervision of medications	3.80
30.	Personal care (bathing, oral hygiene, skin, nail care, shampoo)	2.30
31.	Nutrition (including special diets)	2.79
32.	Fluid balance plus electrolytes	2.90
33.	Elimination	2.68
34.	Rest and sleep	2.71
35.	Physical activity	1.90
36.	Irrigations (including enemas)	2.00
37.	Dressings and bandages	1.55
38.	Formal exercise program	1.74
39.	Rehabilitation (other than formal exercise)	1.17
40.	Prevention of complications and infections	2.14
41.	Recreation, diversion	.50
42.	Clinical procedures - urinalysis, B/P	3.20
43.	Special treatments (e.g., care of colostomy or catheter care)	2.85
44.	Procedures and techniques taught to patient	2.60

Subscale VII: Promotion of Physical and Emotional Health by Direction and Teaching. Only one item (46) produced a mean score above 1.0. The remaining five items all scored below the 1.0 level. The mean scores for Subscale VII are shown in Table 18.

Table 18

Subsample Mean Scores for Items in Subscale VII: Promotion of Physical and Emotional Health by Direction and Teaching

Item #	Item	Mean Score
45.	Plans for medical emergency evident	.31
46.	Emotional support to patient	1.19
47.	Emotional support to family	.26
48.	Teaching promotion and maintenance of health	.19
49.	Evaluation of need for additional resources (e.g., spiritual, social service, homemaker service, physical or occupational therapy)	.41
50.	Action taken in regard to needs identified	.26

Research Questions

The first research question asked: "Are the quality scores of the sample significantly different on the basis of age and sex?" Multiple analysis of variance was

computed for differences in quality scores by age and sex and was not found to be significant at the .05 level. Therefore, it was concluded that the random sample represents a single multivariate population with reference to the variables of age and sex.

Research question two asked: "How predictive of the total scores are the subscale scores of the Phaneuf Nursing Audit?" The scores of the 300 cases were statistically treated by the multiple regression method. The data are presented in Table 19. The F statistic ranged from 341.35 for Subscale II to 442.66 for Subscale III. All seven subscales were found to be significantly predictive (p=.05) of the total scores.

Table 19

F Statistic and Degrees of Freedom for the Total Scores of Subscales I-VII

Subscale	df	F *
I	296	414.59
II	298	341.35
III	297	442.66
IV	294	385.22
V	295	395.73
VI	293	401.77
VII	292	377.94

* p=<.05,n=300

Research question three: "What is the internal consistency of all the items on the Phaneuf Nursing Audit?" Alpha coefficients were computed and reliability estimated internal consistency for the 50-item scale and seven subscales. Data in Table 20 lists the mean Alpha coefficients for the subscales. They ranged from -.08 for Subscale VII to .97 for Subscale IV. Except for Subscale I (r=.27) and Subscale VII (r=-.08), the remaining subscales produced moderately low but acceptable r's of .70 or above.

Subscale I generated the lowest positively scored alpha (.27) of the seven subscales. Item 2 did not generate any

Table 20

Estimates of Internal Consistency by Subscale for the Phaneuf Nursing Audit

Subscale	r
I: Application and Execution of Physician's Legal Orders	.27
II: Observation of Symptoms and Reactions	.73
III: Supervision of the Patient	.70
IV: Supervision of Those Participating in Care (Except the Physician)	.97
V: Reporting and Recording	.73
IV: Application and Execution of Nursing Procedures and Techniques	.73
VII: Promotion of Physical and Emotional Health by Direction and Teaching	-.08

variance and two other items (1 & 3) consistently generated almost perfect scores with small amounts of variance. Three of the six items discriminated very little and consistently generated high scores.

The alpha coefficient for the Phaneuf instrument was 0.85 and is an acceptable estimate of internal consistency. This value accounts for 72% of variance from universe-score difference and 28% from error variance.

The 63-case subsample was inadequate to address statistically research question four: "What comparisons can be made concerning scores generated by the Phaneuf Nursing Audit and by scores resulting from a modified form of the Phaneuf Nursing Audit," and research question five: "What constructs can be derived from a modified form of the Phaneuf Nursing Audit?" The inadequacy was due to the small size of the subsample.

Summary

Three hundred medical records of terminally ill cancer patients were retrospectively audited with the Phaneuf Nursing Audit instrument. The findings indicate that nurses do not appropriately sign, date, nor indicate their caregiver status in the patient's medical record. Documentation is adequate for nursing activities that require tasks or technical duties, but inadequate for the initiation and ongoing assessment of the patient by the nurse. The initiation and updating of nursing care plans,

teaching patients and families, and emotional support of both patients and families were not adequately documented in the patient records. Nurses do not document supervision of others caring for the patient and promotion of physical and emotional health by direction and teaching. Physical activity, recreation, and diversion are also not well documented by nurses. The nurses consistently documented physician-dependent behavior concerning orders, diagnosis, and prompt execution of the medical orders.

The 63-case subsample in general evidenced the same common deficits in documentation concerning the emotional aspects of care and lack of patient/family interaction. The physician-dependent items were the most well documented activities, as were items dealing with the technical aspects of care.

A quality score profile completed for the subscales indicates that Subscale I (Application and Execution of Physician's Legal Orders) was rated in the good category. Subscale II (Observation of Symptoms and Reactions) resulted in the only excellent rating. Subscale III (Supervision of the Patient) and Subscale V (Reporting and Recording) were both in the incomplete range. Subscale VI (Application and Execution of Nursing Procedures and Techniques) also scored in the incomplete range. The lowest ratings were for Subscale IV (Supervision of Those Participating in Care) and Subscale VII (Promotion of

Physical and Emotional Health by Direction and Teaching). Both Subscale IV and Subscale VII resulted in a rating in the range of poor.

The quality score distribution for the sample produced seven excellent, 181 good, 91 incomplete, and 21 poor scores. No unsafe scores were found in the audit.

The sample was representative of the population on the basis of the variables age and sex. All seven subscales were found to be predictive of the total audit scores. An acceptable reliability was found for five of the seven subscales. The overall reliability for the Phaneuf instrument was moderately high. Research questions pertinent to the subsample were unable to be addressed due to the small size of the subsample.

CHAPTER V

CONCLUSIONS, DISCUSSION, IMPLICATIONS, AND RECOMMENDATIONS

In this chapter are presented conclusions and discussion that are based on the findings of the study. Implications and recommendations for nursing service administration, education, practice, and research are presented.

The purpose of this study was to develop estimates of reliability and validity for the Phaneuf Nursing Audit Tool. Three hundred medical records of terminally ill cancer patients who expired during their final hospitalization were audited. In addition, a subsample of 63 medical records was used to generate unweighted scores in order to make comparisons to the original instrument.

Conclusions

Based on the findings the following conclusions are drawn from this study:

1. Nurses providing care for terminally ill cancer patients consistently: (a) carry out physicians' orders, (b) document physical care and technical tasks performed for patients.

2. Nurses providing care for terminally ill cancer patients inconsistently: (a) document initial and ongoing patient assessments, (b) assess the emotional status of patients upon admission and throughout the period of hospitalization, (c) initiate and update nursing care plans as the patient's condition changes, and (d) document activities pertaining to teaching, supervision of others participating in patient care, and promotion of health.

3. Basic nursing documentation such as level of caregiver, signature, and appropriate dates is not consistently found in the patient's medical record.

4. Documentation of the quality of nursing care, as judged by Phaneuf's quality score profile, provided terminally ill cancer patients borders on the lower limits of the good category.

5. The sample appears representative of the population on the basis of distribution by age and sex.

6. There is support of moderate reliability and predictive validity of the Phaneuf instrument.

7. No comparisons could be made, nor constructs identified, between the Phaneuf instrument and a modified form of the Phaneuf instrument.

8. The Phaneuf Nursing Audit instument appears to be adequate for categorizing the quality of nursing care documented in medical records of terminally ill cancer patients.

Discussion

The findings of this research are consistent with other reported audits (Phaneuf, 1972; Hung cited in Phaneuf, 1976) utilizing the Phaneuf instrument. Nurses continue to inappropriately sign and date and fail to indicate their caregiver status. The dependent nurse function of carrying out physicians orders was the one behavior documented most consistently. Nurses need to assume a higher level of responsibility for the development of care plans. The patient's physical condition was documented only one-third of the time. The inadequate documentation of the emotional status is truly an alarming finding considering the patient population of cancer patients who were critically ill and had to cope with the inevitability of approaching death. Ongoing assessment of the patient's emotional status should without exception be an integral part of nursing actions and expected patient care.

Audit criteria concerned with tasks and procedures are consistently executed by nurses. Based on documentation in charts interactions dealing with patient and family were less frequently attended to by the nursing staff. The compilation and use of nurse-generated health histories were not done regularly nor consistently. This lack of documented information is disconcerting since many subsequent nursing actions cannot be intelligently

performed without an adequate data base profiling the patient's psycho-social-physical status. Proper execution of other nurse functions requires this initial and ongoing assessment. Based on documentation in the patient record there is a lack of initiation of a total nursing view of the patient upon admission as well as a continuous deficit throughout the period of hospitalization.

The patient can be characterized as being physically cared for by nurses who carry out the physicians' orders but infrequently delve into the emotional status of the patient and have minimal interaction with the patient and/or family. Donabedian (1980) supports the finding that technical care is usually overemphasized while management of the interpersonal aspects of care tend to be ignored. Findings of Phaneuf's 1968 study also contend that nurses seldom fail to carry out physician dependent functions while rarely recording assessments of the patient's emotional status or interactions with patients and/or family.

Phaneuf (1972) cites this same type of depersonalized nursing care as especially troubling for institutions caring for the critically ill and dying patient. Phaneuf's concern is extremely cogent for this study since these important nursing functions are not consistently being documented for a group of terminally ill cancer patients during their final hospitalization.

Visual scrutiny of Subscale I, entitled Application and Execution of Physician's Legal Orders, reveals differing content of items. The first three items (medical diagnosis complete, orders complete, and orders current) are dependent on the physician and are not nurse dependent. In some institutions nurses or clerical staff may be required to remind physicians if policies are not met concerning their physician responsibilities. The next three items (orders promptly executed, evidence nurse understood cause and effect, and evidence nurse took health history into account) are nurse dependent behaviors that the nurse can carry out independent of the physician. The concern is that items which are physician dependent behaviors are being rated in a nursing scale that measures quality of nursing. Since the documentation for these physician dependent behaviors is present most of the time, a yes or no response could be used to indicate the presence of information necessary for the execution of nursing functions.

Process evaluation based on Donabedian's framework appears to be appropriate for studying elements of nursing care. Study samples based on targeted groups of patients by diagnosis may not be representative of the population of terminally ill cancer patients. The findings of this study are only as valid as the documentation available in the medical records and may account for some of the error

variance, a portion of which may stem from the accurateness and completeness of documentation in the patient's chart. Since criteria for admission to the sample excluded any grossly incomplete medical records, the findigs of the study may be skewed positively towards the upper limits of the total possible scores.

Implications and Recommendations

Implications for Nursing Service Administration.

The findings of this study imply that nursing service administration should provide managerial direction to establish and/or maintain acceptable standards of nursing documentation. Financial resources should be provided by nursing service administration in order for acceptable standards to be attained.

Recommendations for Nursing Service Administration.

Nursing Service Administration needs to:

1. Monitor and/or enforce policies and procedures concerning standards of documenting nursing care for terminally ill cancer patients,

2. Recognize the need and provide the opportunity for nurses to develop skill in assessing and monitoring the health status of terminally ill cancer patients,

3. Provide administrative support for management personnel to monitor nursing actions,

4. The support of appropriate staffing in order for care to be provided for terminally ill cancer patients and

their families and in order that nursing care can be properly documented in the medical record during the cycle of care,

5. Develop and/or enforce nursing care standards that involve family members in the care of the terminally ill cancer patient,

6. Initiate and/or continue an ongoing system of auditing patient records with a valid and reliable tool such as the Phaneuf Nursing Audit instrument,

7. Consider financial recognition for staff that effectively meet standards of care for terminally ill patients, such as a career ladder, and

8. Provide for staff education activities in order to remediate identified audit deficits for nurses caring for terminally ill cancer patients.

Implications for Nursing Education. An implication for nursing education is to incorporate the findings of this study into the curriculum of basic nursing programs. The findings of this study should also be incorporated into the curriculum of graduate nursing programs.

Recommendations for Nursing Education. Recommendations for nursing education are as follows:

1. The inclusion and/or expansion of instruction for students, faculty, and practicing nurses in the documentation of the nursing process utilized to provide care for terminally ill cancer patients.

2. The inclusion and/or expansion of instruction in caring for terminally ill cancer patients to encompass the emotional components of care and involvement of family in the patient's care. Documentation of all aspects of the nursing care process should be included and/or expanded in undergraduate nursing programs.

3. To require that students be able to perform basic nursing documentation in patient charts (e.g., indicate caregiver status, signature, and date).

4. To require that all students be able to assess, update, and document the patient's current health status.

5. That students be encouraged to become involved in process evaluation nursing research and to generate estimates of validity and reliability for all instrumentation.

6. The support for, and development of continuing education for practicing nurses in the community in conjunction with health care agencies to assess and document care, monitor health status, and involve the family in the care of terminally ill cancer patients.

7. The development of auditing skills with a valid and reliable tool such as the Phaneuf instrument.

8. The monitoring of student bahaviors and assessment of the impact of the curriculum by auditing patient charts with a tool such as the Phaneuf Nursing Audit scale.

Implications for Nursing Practice. The findings support the need for nurses to document accurately all nursing care given to patients. Also, nurses should consistently document all nursing care rendered to patients.

Recommendations for nursing practice. Recommendations for nursing practice include the following:

1. Signature, date, and indication of caregiver status should be included in all nursing entries in the patient's medical record.

2. Care plans should be initiated for all patients by nurses upon admission and regularly updated throughout the patient's period of hospitalization.

3. A health history should be compiled and documented by the nurse for each patient at the time of admission.

4. Support of the patient's adaptation to their condition and care should be documented by nurses caring for terminally ill patients.

5. Nurses should document interactions with patients, family, and others.

6. Supervision of those participating in patient care needs to be documented by nurses.

7. Nurses should alert patients and families concerning what to report to physicians.

8. Activities involving exercise, recreation, and levels of activity should be documented by nurses.

9. Documentation should be present for procedures and techniques taught to patients.

10. Promotion of physical and emotional health by direction and teaching should be documented in the patient's medical record.

11. Nurses need to participate in continuing education courses related to caring for the terminally ill.

12. Nurses should participate in the auditing of patient's medical records with a tool such as the Phaneuf Nursing Audit.

13. Nurses need to be involved in the remediation of deficits identified in patient audits, and

14. Nurses should be involved in the peer review of nursing documentation.

Implications for Nursing Research. Further research should be conducted to address such questions as:

1. Are nurses' ages, levels of education, years in practice, and experiences in caring for terminally ill cancer patients contributing factors in the quality of documentation in patient's medical records?

2. Is there a relationship between nurse/patient ratio and quality scores on the Phaneuf Nursing Audit?

3. Is there a difference in quality scores by cancer site?

Recommendations for further research. Recommendations for further research identified from this study include:

1. Replicate this study with a larger sample, across settings, and include other targeted patient population groups for comparison;

2. Conduct longitudinal studies that generate scores over time in order to identify norms and trends of quality scores for specific patient populations, agencies, and regions of the country;

3. Combine outcome and structure research tools with the Phaneuf instrument in order to compare and evaluate a larger portion of phenemona constituting nursing care;

4. Initiate a study with an adequate sample utilizing a modified Phaneuf scale in order to identify alternative groupings of the 50 items based on the statistical loading of items onto factors; and

5. Conduct a follow-up study in the same setting in order to identify changes related to documentation that have occurred since 1979.

Summary

The findings of this study indicate that nurses continue to document the technical aspects of patient care and consistently carry out physicians' orders. However, they do not record basic documentation, initiate and maintain patient health assessments and care plans. The emotional aspects of care involving both patients and their families are areas which are consistently absent in retrospective audits of patients' medical records.

Auditing the nursing care rendered to 300 terminally ill cancer patients during their final hospitalization revealed a lack of documentation for basic components of care expected to be provided any patient population. The Phaneuf Nursing Audit tool demonstrated moderate support for tentative estimates of validity and reliability. If nursing is to develop a scientific base for practice, administration, education, and research, efforts must be directed toward the monitoring of nursing care with valid and reliable tools such as the Phaneuf Nursing Audit.

REFERENCES

American Nurses' Association. Practice standards. Kansas City, Mo.: Author, 1973.

Atwood, J. A research perspective. Nursing Research, 1980, 29, 104-108.

Batey, M. (Ed.). Communicating nursing research: Methodological issues in research (Vol. 3). Boulder, Colo.: Western Interstate Commission for Higher Education, 1970.

Best, J. Research in education (3rd ed.). Englewood Cliffs, N.J.: Prentice-Hall, 1977.

Bloch, D. Evaluation of nursing care in terms of process and outcome: Issues in research and quality assurance. Nursing Research, 1975, 24, 259.

Bloch, D. Interrelated issues in evaluation and evaluation research: A researcher's perspective. American Journal of Nursing, 1980, 29, (2), 69-73.

Brown, F. Principles of educational and psychological testing (2nd ed.). New York: Holt, Rinehart and Winston, 1976.

Campbell, D. Recommendations for APA test standards regarding construct, trait, or discriminant validity. American Psychologist, 1966, 63, 546-553.

Carmines, E., & Zeller, R. Reliability and validity assessment. Beverly Hills: Sage Publications, 1979.

Carter, J., Hilliard, M., Castles, M., Stoll, L., & Cowan, A. Standards of nursing care: A Guide for evaluation (2nd ed.). New York: Springer Publishing, 1976.

Chance, K. The quest for quality: An exploration of attempts to define and measure quality nursing care. Image, 1980, 12, 41-45.

Cozby, P. Behavioral research. Palo Alto, Calif.: Mayfield Publishing, 1977.

Cronbach, L., & Meehl, P. Construct validity in psychological tests. Psychological Bulletin, 1955, 52, 281-302.

DeGeyndt, W. Five approaches for assessing the quality of care. Hospital Administration, 1960, 15, 21-42.

Donabedian, A. Evaluating the quality of medical care. Milbank Memorial Fund Quarterly, (Part 2), 1966, 44, 166-206.

Donabedian, A. Measuring and evaluating hospital and medical care. New York Academy of Medicene, 1976, 52, (1), 51-59.

Donabedian, A. The definition of quality and approaches to its assessment. Ann Arbor, Mich.: Health Administration Press, 1980.

Donabedian, A. The criteria and standards of quality. Ann Arbor, Mich.: Health Administration Press, 1982.

Fox, D. Fundamentals of research in nursing (2nd ed.). New York: Meredith Publishing, 1970.

Froebe, D., & Bain, J. Quality assurance programs and controls in nursing. St. Louis: C. V. Mosby, 1976.

Galton, M., and Reilly, M. Evaluation of nursing care through a nursing audit. Australian Nurses Journal, 1977, 6 (11), 34-36.

Goodwin, L., & Prescott, P. Issues and approaches to estimating interrater reliaility in nursing research. Research in Nursing and Health, 1981, 4, 323-337.

Grier, B. Florence Nightingale and the evaluation of health care. Unpublished Paper. Rush/Presbyterian/St. Luke's Medical Center, Chicago, Ill., 1981.

Hegyvary, S., & Haussman, R. Monitoring nursing care quality. Journal of Nursing Administration, 1975, 5, (5), 17-26.

Horn, B. Establishing valid and reliable criteria: A researcher's perspective. Nursing Research, 1980, 29, 88-90.

Huck, S., Cormier, W., & Bounds, G. Reading statistics and research. Evanston: Harper & Row, Publishers, 1974.

Isler, C. Nursing audit: New yardstick for patient care. Registered Nurse, 1974, 37, (12), 31.

Jelinek, R., & Dennis, A. A review and evaluation of nursing productivity (Publication No. [HRA] 77-15). Bethesda, Md.: U. S. Department of Health, Education, and Welfare, 1976.

Joint Commission on Accreditation of Hospitals. Accreditation manual for hospitals. Chicago: Author, 1970.

Kerlinger, F. Foundations of behavioral research (2nd ed.). New York: Holt, Rinehart & Winston, Inc., 1973.

Kerlinger, F., & Pedhazur, E. Multiple regression in behavioral research. Chicago: Holt, Rinehart & Winston, Inc., 1973.

King, E. Determining and interpreting test validity, reliability and practicality. Nurse Educator, 1979, 5, (8), 6-11.

Krueger, J., Nelson, A., & Wolanin, M. Nursing research: Development, collaboration, and utilization. Germantown, Md.: Aspen Systems, 1978.

Lesnik, M., & Anderson, B. Nursing Practice and the Law (2nd ed.). Philadelphia: J. B. Lippincott, 1955.

Lindeman, C. Nursing intervention with the presurgical patient: Effectiveness and efficiency of group and individual teaching, phase two. Nursing Research, 1972, 21, 196-209.

Minium, E. Statistical reasoning in psychology and education. New York: John Wiley, 1970.

Nunnally, J. Psychometric theory. New York: McGraw Hill, 1967.

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

Nunnally, J., & Durham, R. Validity, reliaility, and special problems of measurement. In E. Struening & M. Guttentag (Eds.), Handbook of Evaluation Research (Vol. 1). Beverly Hills: Sage Publications, 1975.

Phaneuf, M. A nursing audit method. Nursing Outlook, 1964, 11, 42-45.

Phaneuf, M. Nursing audit for evaluation of patient care. Nursing Outlook, 1966, 14, 51-54.

Phaneuf, M. Analysis of a nursing audit. Nursing Outlook, 1968, 16, 57-60.

Phaneuf, M. Quality of care: Problems of management. American Journal of Public Health, 1969, 59, 1327-1832.

Phaneuf, M. The nursing audit profile for excellence. New York: Appleton-Century-Crofts, 1972.

Phaneuf, M. The nursing audit: Self regulation in nursing practice (2nd ed.). New York: Appleton-Century-Crofts, 1976.

Polit, D., & Hungler, B. Nursing research: Principles and methods. New York: J. B. Lippincott, 1978.

Staropoli, C., & Waltz, C. Developing and evaluating educational programs for health care providers. Philadelphia: F. A. Davis, 1970.

Statistical package for the social sciences X. New York: McGraw-Hill Book Company, 1983.

Ventura, M., & Hageman, P. Testing of the reliability, validity, and sensivitity of quality of nursing care measures: Final report. Buffalo, N.Y.: Veterans Administration Medical Center, 1978.

Ventura, M. Correlation between the quality patient care scale and the Phaneuf Audit. International Journal of Nursing Studies, 1980, 3, (17), 155-162.

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

Wandelt, M., & Stewart, D. Slater nursing competencies rating scale. New York: Appleton-Century-Crofts, 1975.

Winer, B. Statistical principles in experimental design. New York: McGraw-Hill, 1971.

REFERENCE NOTES

1. Phaneuf, M. Personal communication, July 16, 1980.

Appendix A

The Phaneuf Nursing Audit

PART I. HOSPITAL OR NURSING HOME AUDIT

Data must be held in STRICT confidence and MUST NOT BE FILED with patient's record.

All Entries To Be Completed By Trained Clerk

1. Name of patient: (LAST) (FIRST) 2. Sex 3. Age 4. Date admitted 5. Discharge date

6. Name of institution: 7. Floor 8. Medical supervision Private ☐ Ward ☐ OPD/Clinic ☐

9. Complete diagnosis(es):

10. Admitted by referral from: Physician on staff ☐ M.D. not hospital affiliated ☐ Clinic/OPD ☐ 11. Via emergency ☐

12. Patient discharged to: Self-care ☐ Family care ☐ PHN Agency ☐ Other specify: ☐ Died ☐ Unknown ☐

13. If patient died: M.D. present ☐ M.D. promptly notified ☐ Family present ☐ Family promptly notified ☐ 14. If patient Catholic: Last rites given: YES ☐ NO ☐

15. All nursing entries signed by name and dated: YES ☐ NO ☐ 16. Nursing entries show whether made by professional, practical, student nurse, or other: YES ☐ NO ☐

17. Patients' clothing, valuables, and other personal items were accounted for in accordance with policy: YES ☐ NO ☐

	YES	NO
18. Operative and other patient or family consent forms completed as required by policy	___	___
19. A. Were there any accidents or other special incidents?	___	___
B. If yes, chart indicates report was submitted to administration	___	___
C. Or, report is part of chart	___	___
20. A. Kardex in use	___	___
B. If yes, Kardex becomes part of permanent chart	___	___
21. Nursing care plan is recorded in the chart	___	___
22. A. Nursing admission entry shows assessment of patient's condition: physical	___	___
emotional	___	___
B. Nursing discharge entry shows assessment of patient's condition: physical	___	___
emotional	___	___

PART II. NURSING AUDIT CHART REVIEW SCHEDULE

All Entries To Be Completed By A Member Of the Nursing Audit Committee

(Please check in box of choice; DO NOT obscure number in box.)

Name of patient: _____
(LAST) (FIRST)

	YES	NO	UNCERTAIN	TOTALS
I. APPLICATION AND EXECUTION OF PHYSICIAN'S LEGAL ORDERS				
1. Medical diagnosis complete	7	0	3	
2. Orders complete	7	0	3	
3. Orders current	7	0	3	
4. Orders promptly executed	7	0	3	
5. Evidence that nurse understood cause and effect	7	0	3	
6. Evidence that nurse took health history into account	7	0	3	
(42) TOTALS		0		☐
II. OBSERVATION OF SYMPTOMS AND REACTIONS				
7. Related to course of above disease(s) in general	7	0	3	
8. Related to course of above disease(s) in patient	7	0	3	
9. Related to complications due to therapy (each medication and each procedure)	7	0	3	
10. Vital signs	7	0	3	
11. Patient to his condition	7	0	3	
12. Patient to his course of disease(s)	5	0	2	
(40) TOTALS		0		☐
III. SUPERVISION OF THE PATIENT				
13. Evidence that initial nursing diagnosis was made	4	0	1	
14. Safety of patient	4	0	1	
15. Security of patient	4	0	1	
16. Adaptation (support of patient in reaction to condition and care)	4	0	1	
17. Continuing assessment of patient's condition and capacity	4	0	1	
18. Nursing plans changed in accordance with assessment	4	0	1	
19. Interaction with family and with others considered	4	0	1	
(28) TOTALS		0		☐
IV. SUPERVISION OF THOSE PARTICIPATING IN CARE (EXCEPT THE PHYSICIAN)				
20. Care taught to patient, family, or others, nursing personnel	5	0	2	
21. Physical, emotional, mental capacity to learn considered	5	0	2	
22. Continuity of supervision to those taught	5	0	2	
23. Support of those giving care	5	0	2	
(20) TOTALS		0		☐
V. REPORTING AND RECORDING				
24. Facts on which further care depended were recorded	4	0	1	
25. Essential facts reported to physician	4	0	1	
26. Reporting of facts included evaluation thereof	4	0	1	
27. Patient or family alerted as to what to report to physician	4	0	1	
28. Record permitted continuity of intramural and extramural care	4	0	1	
(20) TOTALS		0		☐

PART II. NURSING AUDIT CHART REVIEW SCHEDULE (cont.)

VI. APPLICATION AND EXECUTION OF NURSING PROCEDURES AND TECHNIQUES	YES	NO	UNCERTAIN	TOTALS	DOES NOT APPLY
29. Administration and/or supervision of medications	2	0	0.5		2
30. Personal care (bathing, oral hygiene, skin, nail care, shampoo)	2	0	0.5		2
31. Nutrition (including special diets)	2	0	0.5		2
32. Fluid balance plus electrolytes	2	0	0.5		2
33. Elimination	2	0	0.5		2
34. Rest and sleep	2	0	0.5		2
35. Physical activity	2	0	0.5		2
36. Irrigations (including enemas)	2	0	0.5		2
37. Dressings and bandages	2	0	0.5		2
38. Formal exercise program	2	0	0.5		2
39. Rehabilitation (other than formal exercise)	2	0	0.5		2
40. Prevention of complications and infections	2	0	0.5		2
41. Recreation, diversion	2	0	0.5		2
42. Clinical procedures - urinalysis, B/P	2	0	0.5		2
43. Special treatments (e.g., care of tracheotomy, use of oxygen, colostomy or catheter care, etc.)	2	0	0.5		2
44. Procedures and techniques taught to patient	2	0	0.5		2
(32) TOTALS		0			
VII. PROMOTION OF PHYSICAL AND EMOTIONAL HEALTH BY DIRECTION AND TEACHING					
45. Plans for medical emergency evident	3	0	1		3
46. Emotional support to patient	3	0	1		3
47. Emotional support to family	3	0	1		3
48. Teaching promotion and maintenance of health	3	0	1		3
49. Evaluation of need for additional resources (e.g., spiritual, social service, homemaker service, physical or occupational therapy)	3	0	1		3
50. Action taken in regard to needs identified	3	0	1		3
(18) TOTALS		0			
			TOTAL SCORE		
			FINAL SCORE		

Appendix B

Permission for Use of the Instrument

This is to certify that Robert W. Vogler is granted permission to use the Phaneuf Nursing Audit for purposes of dissertation research in a study to generate estimates of validity and reliability. The instrument appears in The Nursing Audit Self-Regulation in Nursing Practice, by Maria Phaneuf, 1976 and may be revised and/or edited as necessary for this specific research effort.

Raymond Lamm Peters Editor

Signature

June 21, 1980

Title

Appendix C

Criteria for Subscales I-VII

EXPLANATION OF AUDIT SCHEDULE COMPONENTS

FUNCTIONS AND SUBCOMPONENTS

Function I. Application and Execution of Physician's Legal Orders

1. *Medical Diagnosis Complete.* The diagnosis is clear enough to permit intelligent execution of the nursing functions. A diagnosis which conforms in terminology with that of the International Classification of Diseases, published by the U.S. Department of Health, Education and Welfare, ordinarily suffices.

   At varying points in patient care, as when clinically unexplained changes occur, the maximum nursing base may be a tentative clinical diagnosis or other significant data that justify intervention.

   Where patients have multiple diagnoses, the same rules apply. Here, one of the hazards in nursing and for the patient, however, is that the primary diagnosis in relation to which care is being given remains the sole focus of care, whereas the other diseases or disorders involved may be equally important in the nursing process. The patient who had cholecystitis with cholecystectomy, and also has longstanding diabetes mellitus, may receive much nursing attention for the surgical problem but only cursory attention to his problems with diabetes.

2. *Orders Complete.* The physician's orders are clear, explicit, and conclusive *when looked at in regard to the patient, as well as to the diagnosis and other clinical data.* Orders for medications should include the dosage and frequency of administration, and the route of administration (unless it is clear from the nature of the medication, as for aspirin or insulin). Orders should be specific not only when medications are administered by the

nurse, but also when self-administered by the patient or given by family members or other responsible persons.

3. *Orders Current.* Orders are up to date according to pertinent institutional or agency policy *and* nursing judgment. For example, an order for Seconal may fall within the stop-order policy limit. But if there is evidence that Seconal is causing untoward effects in the patient, the nurse will withhold the medication and consult the physician about the situation rather than adhere to policy only.
4. *Orders Promptly Executed.* The chart shows reasonable and appropriate timing between the giving of the order and compliance with it. There should be adherence to institutional or agency policy in regard to the dating of the orders, the recording of the time at which the orders are written, and the recording of the date and time of execution of the orders.
5. *Evidence That the Nurse Understood Cause and Effect.* The chart shows that the nurse knew what she was doing and why she was doing it. A nurse performing any service ordered by the physician is legally obligated to understand the cause and effect of that service before performing it. The nurse is required to understand not only the basis for and anticipated therapeutic results of performance, but also the possible side effects of other complications. It cannot be too strongly emphasized that the nurse's right to perform any function is absolutely contingent upon her ability to understand its underlying reason and its anticipated effect, as well as upon her ability to perform the function.
6. *Evidence That the Nurse Took the Health History Into Account.* The chart reflects recognition that knowledge of pertinent points in the patient's past pattern of health and illness are vital to intelligent current nursing care. The purpose of the history is to develop data from which to make nursing assessments of strengths, weaknesses, and life style which are taken into account when planning nursing intervention relative to health-illness problems.

Function II. Observations of Symptoms and Reactions

7. *Related to the Course of the Above Disease in General.* There is evidence that the nurse understands the disease in the textbook or classic sense and is observing the patient with the classic picture as her clinical frame of reference.

   By this is meant that the natural history of the disease from which the patient suffers should be known by the nurse and used as the clinical base for developing the nursing process. In this regard, the paradigm developed by Leavell and Clark is useful because it depicts the pattern of movement through prepathogenesis, early pathogenesis, discernible early disease, advanced disease, and convalescence, with possible outcomes of recovery, a chronic state, disability, or death.
8. *Related to the Course of the Above Disease in This Patient.* There is evidence that, in addition to the knowledge of the disease in item 7, there are observations of the patient's individual response to the disease and its treatment.

   This simply means progression from consideration of the natural history of the disease in any man and the natural history of the disease in the par-

ticular man, which may be influenced by his heredity, his general health, and his life situation, as well as by his treatment and response to treatment.

9. *Related Complications Due to Therapy (Each Medication and Each Treatment).* Recorded observations relate to expected therapeutic and possible or unexpected untoward side effects.

   The observations are one reflection of the nurse's ability to understand cause and effect relationships in nursing management.

10. *Vital Signs.* When indicated by the patient's situation, recording includes: temperature; quality of pulse, as well as rhythm and rate; quality of respirations, as well as rate; blood pressure; tone, temperature, and color of the skin; and observations pertinent to feeling tone—that is, the patient's affective state.

    Here the emphasis is on collection of data so that patterns and trends in the vital signs are clear. The recording of a single vital sign, such as one blood pressure determination, is meaningless because it is the trend in blood pressures that indicates need for, and response to, therapy.

11. *Patient to His Condition.* There is evidence that attention was given to the patient's attitude toward his clinical condition and life situation as it influences and is influenced by the clinical condition.

    "Attention" means careful consideration of behaviors reflective of attitude. This includes use of direct, indirect, and reflective questions to the patient aimed at eliciting attitudinal responses, as well as observation of nonverbal behavior.

12. *Patient to His Course of Disease.* There is evidence that attention was given to the demonstrable degree of the patient's understanding and acceptance, rejection, or ambivalence toward his specific disease and illness.

    Attention here is literally twofold: attention to the disease that is to the pathological process; and attention to the illness that is the acute or chronic manifestations of the pathologic process. For example, it is possible for a patient to reject his disease but to accept the illness it causes, or to accept the disease but reject his illness. Nursing intervention carried out without recognition of the patient's position will fall short of its mark.

Function III. Supervision of the Patient

13. *Evidence That Initial Nursing Diagnosis Was Made.* The chart shows that nursing problems were determined and categorized as the basis for nursing care plans directed toward solution of the problems. This diagnosis should be made as soon as possible after the first nursing contact with the patient. In some charts, nursing care plans strongly suggest that a diagnosis was made. In this event, evidence of the implicit diagnosis should be taken into account.

    Since "to determine" means to establish after consideration, investigation, or calculation, and "to categorize" means to classify into specified divisions, it is obvious that initial "nursing diagnosis" as here used encompasses the steps in the nursing process up to the point of formulation of the plan of care.

14. *Safety of the Patient.* There is recorded evidence of precautions taken to prevent physical injury.

These precautions include assistance in early ambulation and other activities involving neuromuscular functions which are difficult for the patient and encompass environmental safeguards as well.

15. *Security of the Patient.* There is evidence of work that helps in maintaining a therapeutic environment for the patient.

    This work includes support of productive interpersonal relationships, as well as attention to the physical setting in which the human interactions occur.

16. *Adaptation (Support of Patient in Reactions to Condition and Care).* There is evidence of attempts to help the patient adjust to his changing condition, to the course of his illness, to his care, and to his anticipated future.

    These attempts include helping the patient to accept attainable therapeutic goals; helping reduce the patient's anxiety, fear, and doubt; helping him toward self-confidence and confidence in his care; and helping the patient to exert the physical and emotional efforts required in his situation, in accordance with his capacities.

17. *Continuing Assessment of Patient's Condition and Capacity.* The chart reflects ongoing evaluation of the current status and situation of the patient and the effects of care, with analysis of current nursing problems.

    This continuing assessment involves both the collection of data with validation of them and interpretation of that data with validation of the interpretation as a base for modification or revision of the plan of care.

18. *Nursing Care Plans Changed in Accordance with Assessment.* There is evidence that the plan of care was adapted as nursing problems were altered by changes in the patient's condition and capacity.

    In relation to assessment, the difference between this and item 17 is that item 17 emphasized continuity of assessment, as opposed to assessment of one or another single aspect of the patient's condition or capacity.

    This subcomponent, however, is primarily addressed to the question of whether the nursing care plan was appropriately altered as the patient's condition and capacity changed.

19. *Interaction with the Family and with Other People Considered.* There is evidence of concern for the people in contact with the patient, with a view toward promoting interactions that are beneficial to all concerned.

    This means that the patient's interactions with his family, his physician, and other people important to them are observed with respect to the interests and concerns reflected therein by them, and use of those observations to advance mutually constructive relationships.

Function IV. Supervision of Those Participating in Care (Except the Physician)

20. *Care Taught to Patient, Family, or Others Participating in His Care.* The chart reflects what care was taught, what guidance and support were given, to whom, and by whom accomplished.

    The care taught includes all activities resumed or assumed by the patient and all the tasks performed by others involved in his care. It is assumed that care has not been taught until the behavior of those taught shows or suggests that learning has occurred.

21. *Physical, Mental, and Emotional Capacity to Learn Considered.* The evidence shows that the ability and readiness of those to be taught, guided, and supported were taken into account.

    Consideration of the learner's capacity includes initial and continuing assessments of the need for and the appropriateness of that which is to be taught, in relation to the ability and the readiness of those being taught.

22. *Continuity of Supervision to Those Taught.* The evidence shows that the results of initial and additional teaching were assessed with appropriate follow-up.

    This subcomponent is also based on the assumption of item 20 that teaching has not occurred until it is reflected in the behavior of the learners, and that activities and tasks in self-care or care given by others are not ordinarily learned through a single exposure to "teaching." The emphasis here is on the follow-up.

23. *Support of Those Giving Care.* The chart reflects the giving of emotional and physical help to those taught and supervised.

    Here the emphasis is on continuing assessment of the ability and readiness of those taught, with appropriate action in accordance with the assessment.

Function V. Reporting and Recording

24. *Facts on Which Further Care Depended Were Recorded.* The information recorded facilitated continuing physician and nurse management of clinical care.

    Minimum information includes observations of symptoms and reactions; evidence of the execution of physician's orders; and data developed as part of the supervision of the patient.

25. *Essential Facts Reported to the Physician.* The chart shows that basic necessary information was conveyed to the physician either in writing or verbally. The facts may be major or minor; it is their importance to the physician and his management of the patient's care that makes them essential or nonessential.

    Essential facts are those indispensible to patient-centered care, as well as those that are clinically significant as discrete facts.

26. *Reporting of Facts Included Evaluation Thereof.* There is evidence that, in reporting facts, nursing judgment concerning their significance or possible importance is included.

    In other words, the emphasis here is on nursing expression of the reason why the facts were considered indispensible to the physician in the management of his patient.

27. *Patient and Family Alerted as to What to Report to the Physician.* There is evidence that patient or family members are directed to report to the physician those factors, signs, symptoms, or situations the direct reporting of which is conducive to patient and family rapport with the physician, or is otherwise mutually advantageous.

    The intent here is to foster communications with the physician about questions which the nurse cannot properly answer, or questions the answering of which by the physician serves a special purpose in the management of

medical or nursing care. One special purpose would be to have anxieties and fears allayed by the physician when he can best accomplish this.

Having the patient and family members report to the physician does not relieve the nurse of responsibility for direct reporting to the physician. In critical matters, the physician may be assisted by receiving separate communications from the patient or family and from the nurse.

28. *The Chart Permitted Continuity of Care.* The chart permits an uninterrupted sequence of care from nurse to nurse, from nurse to physician, and from nurse to other professionals. It is of major importance that the chart indicate succinctly that information vital to the patient's therapeutic regimen was reported to the physician.

    The question to be answered in this subcomponent is not whether there actually was continuity in care, but whether continuity of care was possible with use of the information on the chart.

Function VI. Application and Execution of Nursing Procedures and Techniques

29. *Administration of Medications/Supervision of Their Use.* Whether medications are given by the nurse, or whether the nurse is supervising the patient or the family in the taking or giving, the chart reflects nurse or patient and family awareness of expected therapeutic results and possible untoward side effects.

    For every medication, including those administered by the physician, there are anticipated therapeutic effects and possible untoward side effects, including reactions of intolerance and idiosyncrasy. Wherever more than one medication is used, the possibility of drug incompatibility must also be considered.

30. *Personal Care (Bathing, Oral Hygiene, Skin, Nail, and Hair Care).* The chart indicates appropriate attention to personal care whether the care activities are performed by the patient, a family member, or another person.

    Appropriate attention includes not only concern with cleanliness, but also with grooming conducive to feelings of well-being, personal worth, and dignity.

31. *Nutrition, Including Special Diets.* There is evidence of attention to adequate nutrition as appropriate to the patient's condition, course, and stage of growth and development. If a special diet is used, there is evidence as to whether or not, and to what extent, the diet and the main reasons for it appear to be understood and accepted by the patient and his family.

    Appraisal of the patient's usual eating habits is a part of nursing assessment, whether the patient is on a regular or a special diet. Results are used in formulating and implementing nursing care plans.

32. *Fluid and Electrolyte Balance.* The chart reflects consideration of possible disturbances in body fluid and electrolyte balance, as indicated by the patient's age, condition, and course of illness.

    Considerations of possible disturbances include attention to fluid intake and urinary output; changes in respiratory rate and depth; changes in skin turgor; dryness of skin and mucous membranes; changes in behavior, such as increasing apathy or restlessness; thirst; ascites; and edema.

33. *Elimination.* Evidence that bowel function is considered.

The emphasis should be on what is normal for the patient in health and the deviations that occur because of his illness. Action should follow the patient's pattern as closely as permitted by his condition.

34. *Rest and Sleep.* Evidence that the patient's usual and unusual patterns of rest and sleep are taken into account in planning his regimen and supervising his care.

    Appraisal of the patterns of rest and sleep is also a part of nursing assessment. It permits planning of the regimen with regard to rest and sleep, to follow as closely as possible the patient's natural rhythms. If the usual patterns obviously yield deficits in rest and sleep, the regimen should be planned with the aim of bringing about appropriate alterations of the pattern.

35. *Physical Activity.* The chart shows the relationship between the activity in which the patient engages and the activity which is clinically permissible. Where excess or deficit is found, efforts are made to reconcile actual physical activity with clinically estimated physical tolerance.

    Development of a balance between too much and too little activity requires that the patient understand and accept the reasons that underlie restriction or increase in activity.

36. *Irrigations of Wounds, Canals, Cavities.* Evidence that irrigations are performed as ordered; the results; and, if dressings are used, what kind and whether sterile or clean.

    This subcomponent refers to all types of irrigation and includes enemas. If problems in performance of the procedure arise either in relation to the irrigation or in relation to patient reaction to it, they should be recorded.

37. *Dressings and Bandages.* Evidence that these are applied as ordered or as indicated. Topical applications, if any, should be identified; the kind of dressing used and whether it was sterile or clean should be noted.

    Observations of the wound site and adjacent tissues should be recorded in a manner that permits continuing appraisal of progress in healing and early detection of complications.

38. *Formal Exercise Program.* Indication that a treatment plan is carried out as ordered by the physician or as outlined by a physical therapist at the physician's request.

    Here, the nurse is responsible for seeing that the program is carried on and also that supportive encouragement and assistance are given to the patient.

39. *Rehabilitation (Other Than Formal Exercises).* Evidence of teaching or encouragement toward independent living—range of motion (ROM), active and passive exercises, activities of daily living (ADL), use of aids in ADL. If nursing rehabilitation is not required, there is evidence that the nursing care approach is restorative in nature.

    Activities of daily living require motivation and participation in decision making which leads to the activities, as well as to the ability to perform them. Evaluation of the performance may increase or decrease motivation. At best, encouragement from the nurse reflects knowledge and understanding of this. The restorative approach has the same foundation. Activities of daily living include not only self-care but also other activities which give a positive meaning to the day for the patient.

40. *Prevention of Complications, Including Infections.* Evidence of work toward

maintenance of hygiene, early detection of primary or secondary infections or other untoward symptoms; early detection of complications due to therapy; and prevention of avoidable disabilities, such as contractures.

Consideration of complications that might reasonably be expected, or prevented, is a part of initial and ongoing nursing assessment.

41. *Recreation and Diversion.* The chart indicates specific attention to the patient's need for activities that interest and amuse him and divert his attention from disease and illness.

For counteracting disease- and illness-oriented tasks, activities, and limitations, the importance of recreation and diversion, however simple the related activities may be, cannot be overemphasized as part of orientation toward health.

42. *Clinical Procedures.* The chart shows results of urinalyses and other examinations done by nurses, if any; blood pressure determinations; and results of performance of other general nursing procedures.

43. *Special Treatments, Including Tracheostomy Management, Use of Oxygen, Colostomy Care, Gastric Feedings, Care of Decubiti, etc.* Evidence that the treatments were performed, indication of results, and evaluation thereof; observations pertinent to patient's physical and emotional reactions.

The preparation of the patient for the special treatment is a part of the performance of special treatments. Patient's preferences as to the way in which the procedure is to be performed should be recognized and adhered to, as well as possible. Where it is not safe to follow his preference, the record should indicate efforts to explain this and to enlist his cooperation.

44. *Procedures and Techniques Taught to Patient.* Evidence that any procedure or technique the patient can learn to carry out to his advantage is in fact taught.

Function VII. Promotion of Physical and Emotional Health By Direction and Teaching

45. *Plans for Medical Emergency.* Evidence that, by policy or by specific teaching, patient, family, and other personnel know what to do in situations which are acutely worrisome or dangerous for the patient and which arouse anxiety or fear in those responsible for his care.

Planning for medical emergencies is contingent on assessment of the emergencies that might reasonably be expected to arise, in terms of what the patient and his family perceive as constituting an emergency and what is clinically perceived as an emergeny situation.

In accredited hospitals and nursing homes, there are specific policies for the management of major emergency situations. In auditing, it is necessary to note whether policies were carried out as necessary.

In public health nursing agencies, plans about what to do if medical emergencies arise are developed with patient and family and in conjunction with the physician, if the patient is under private medical care. If the patient is not under private care, the patient and his family should know precisely what clinical facility to use in an emergency, and how to use it.

46. *Emotional Support for the Patient.* Evidence of work toward helping the patient understand and accept his feelings about himself, his condition, and his care, and helping him develop his coping abilities and other potentials.

Provision of emotional support requires assessment of the special needs of the patient, his characteristic behaviors, and his psychosocial and cultural matrix. Without this data base, it is unlikely that the rapport and open communication necessary for providing emotional support will be achieved.

47. *Emotional Support for the Family.* The chart reflects impressions and facts about family reactions toward the patient and his condition, which can be used to help the family accept the patient's condition and their own feelings about it.

Providing emotional support for the family requires the same data base as that required in emotional support of the patient.

48. *Teaching Preventive Health Care.* Evidence of promoting and protecting the health of the patient and his family, and of teaching about secondary prevention, that is, teaching the early detection of signs and symptoms which may indicate new disorders or complications due to established disease.

Assessment of the goals and motivation of the patient and family precedes discussion and use of selected educational tools with them. The teaching plan will be unique for the patient and his family, and its effectiveness dependent on the rapport already established. Minimally, patient and family need to understand the medical and nursing regimen and to understand, accept, and carry out the necessary procedures and activities.

49. *Evaluation of the Need for Additional Resources, Including Spiritual Guidance, Social Services, Occupational Therapy, or Continuity of Nursing Care Under Another Aegis; Homemaker Service.* Evidence that, when indicated, possible needs for consultation or direct service were assessed.

Evaluation of the need for use of additional resources should occur periodically throughout the time the patient is under care. Continuation care planning should be done well in advance of the patient's discharge.

50. *Action Taken in Regard to Needs Identified.* Evidence that nursing action was taken, with the knowledge of the patient's attending physician, for needs identified as relating to the promotion, by direction and teaching, of the patient's physical and emotional health.

It is of course useless to identify and categorize needs and problems unless action is taken directly or indirectly to help meet the needs and to solve or alleviate the problems.

Appendix D

Does Not Apply Guide

"Does Not Apply" Value Guide

Total of "Does Not Apply" Items	"Does Not Apply" Score Value
0	1.00
2	1.01
3 or 4	1.02
5 or 6	1.03
7 or 8	1.04
9 or 10	1.05
11 or 12	1.06
13 or 14	1.07
15	1.08
16 or 17	1.09
18 or 19	1.10
20	1.11
21 or 22	1.12
23	1.13
24 or 25	1.14
26	1.15
27 or 28	1.16
29	1.17
30 or 31	1.18
32	1.19
33 or 34	1.20
35	1.21
36	1.22
37 or 38	1.23
39	1.24
40	1.25
41	1.26
42 or 43	1.27
44	1.28
45	1.29
46	1.30
47	1.31
48 or 49	1.32
50	1.33

Appendix E

Modified Phaneuf Nursing Audit Scoring Scale

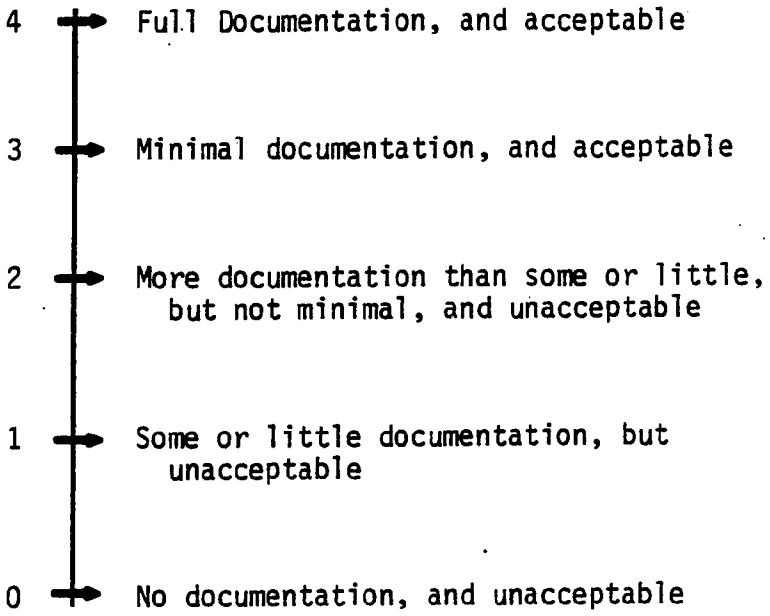

Figure 2. Modified Phaneuf Nursing Audit Scoring Scale

Appendix F

Institutional Review Board Approval

FORM 4

UNIVERSITY OF ALABAMA IN BIRMINGHAM
IDENTIFICATION AND CERTIFICATION OF APPLICATIONS FOR SUPPORT OF RESEARCH INVOLVING HUMAN SUBJECTS

THIS FORM IS REQUIRED FOR ALL APPLICATIONS FOR RESEARCH AND RESEARCH TRAINING GRANTS, PROGRAM PROJECTS AND CENTER GRANTS, DEMONSTRATION GRANTS, FELLOWSHIPS, TRAINEESHIPS, AWARDS, AND OTHER PROPOSALS WHICH MIGHT INVOLVE THE USE OF HUMAN SUBJECTS, REGARDLESS OF ANY SOURCE OF FUNDING. THIS FORM IS NOT APPLICABLE TO APPLICATIONS FOR GRANTS LIMITED TO THE SUPPORT OF CONSTRUCTION, ALTERATIONS AND RENOVATIONS, OR RESEARCH RESOURCES.

THE PUBLIC HEALTH SERVICE REQUIRES ASSURANCE FROM GRANTEE INSTITUTIONS SPONSORING RESEARCH, INVESTIGATION AND CARE WHICH MAY INVOLVE HUMAN SUBJECTS THAT THEY WILL CARRY OUT REVIEW OF ALL SUCH PROJECTS IN AGREEMENT WITH THE POLICY AND INSTRUCTION PROVIDED IN "THE INSTITUTIONAL GUIDE TO DHEW POLICY ON PROTECTION OF HUMAN SUBJECTS", MARCH 13, 1975, AND PART 46 OF TITLE 45 OF THE CODE OF FEDERAL REGULATIONS, AS AMENDED. THE PUBLIC HEALTH SERVICE ALSO REQUIRES CERTIFICATION OF THE PERFORMANCE OF THE INITIAL REVIEW. THIS REVIEW SHOULD BE COMPLETED PRIOR TO SUBMISSION OF RELEVANT PROPOSALS, BUT IN ANY EVENT IT MUST BE DONE PRIOR TO AWARD OF FUNDS.

WITH SPECIFIC REFERENCE TO THE PROJECT NUMBERED _____

ENTITLED AN ASSESSMENT OF THE VALIDITY AND RELIABILITY OF PHANEUF NURSING ADUIT SCORES

UNDER THE DIRECTION OF ROBERT W. VOGLER

[ ] 1. THIS APPLICATION DOES NOT INCLUDE RESEARCH INVOLVING HUMAN SUBJECTS, AS DEFINED IN THE FEDERAL REGISTER, MARCH 13, 1975 AND PART 46 OF TITLE 45 OF THE FEDERAL REGULATIONS, AS AMENDED.

[X] 2. THIS APPLICATION INCLUDES RESEARCH INVOLVING HUMAN SUBJECTS. OUR INSTITUTIONAL REVIEW BOARD HAS REVIEWED AND APPROVED IT ON 9-10-80, IN ACCORDANCE WITH OUR ASSURANCE APPROVED BY THE PUBLIC HEALTH SERVICE. THE PROJECT WILL BE SUBJECT TO CONTINUING REVIEW AS PROVIDED FOR IN THAT ASSURANCE.

[ ] 3. NOT AT RISK THIS APPLICATION, WHICH MAY INCLUDE RESEARCH INVOLVING HUMAN SUBJECTS, IS PENDING REVIEW BY THE INSTITUTIONAL REVIEW BOARD AS PROVIDED BY OUR ASSURANCE. CERTIFICATION OF COMPLETION OF THE REVIEW WILL BE PROVIDED AS SOON AS POSSIBLE. (THIS CERTIFICATION WILL BE IN THE FORM INDICATED IN 2. ABOVE. IDENTIFYING THE APPLICATION BY TITLE, INVESTIGATOR, AND, IF KNOWN, BY NUMBER.

THOMAS H. ALPHIN, M.D.
CHAIRMAN, INSTITUTIONAL REVIEW BOARD
UNIVERSITY OF ALABAMA IN BIRMINGHAM

9-10-80
DATE

REVISED 1/79

GRADUATE SCHOOL
UNIVERSITY OF ALABAMA IN BIRMINGHAM
DISSERTATION APPROVAL FORM

Name of Candidate: Robert W. Vogler

Major Subject: Adult Health Nursing

Title of Dissertation: An Assessment of the Validity and Reliability of Phaneuf Nursing Audit Scores

Dissertation Committee:

____________, Chairman

Director of Graduate Program ____________

Dean, UAB Graduate School ____________

Date 11/30/84

PS-1428