A methodology for validating nursing diagnoses

A methodologic approach consisting of four phases is proposed for validating nursing diagnoses. The sequential phases include use of the Delphi technique, magnitude estimation scaling, patient observation, and testing nursing interventions through clinical trials. The proposed methodologic approach is a strategy for researching the diagnosis and treatment of human responses to actual or potential health problems—the foundation for nursing practice.

Joan Grant, DSN, RN, CS
Assistant Professor

Marguerite Kinney, DNSc, RN, FAAN
Professor
School of Nursing
University of Alabama at Birmingham
UAB Station
Birmingham, Alabama

Cathie Guzzetta, PhD, RN, CCRN, FAAN
Codirector
Holistic Nursing Consultants
Bethesda, Maryland

THE NURSING PROCESS is the foundation for nursing practice and has provided the conceptual and clinical perspective for education and practice for more than two decades.[1,2] Aspinall,[1] however, compared the nursing process to a chain. Just as a chain is no stronger than its weakest link, the strength of the nursing process is derived from each of its component parts. Nursing diagnosis may well be the weakest link in the nursing process.

According to Jacobs and Huether,[3] the order of development for a clinical science is identification and description of phenomena of concern, explanation and substantiation of relevant concepts, and validation of nursing interventions and outcome criteria. Unfortunately, while focusing on the other phases of the nursing process, nurses have failed to address nursing diagnoses satisfactorily.[4]

Formal work on specifying and describing phenomena of concern to nursing began in 1973 when the First National Conference on Classification of Nursing Diagnoses con-

Adv Nurs Sci 1990;12(3):65–74

vened to identify and to classify nursing diagnoses. However, a diagnostic validity gap exists between diagnoses approved for testing and the diagnoses actually used in clinical practice, because tentative diagnoses are dependent on empirical validation before they can be accepted for use in clinical practice.[5]

Although a few models have been proposed to validate nursing diagnoses, validation research is still sparse due, in part, to the lack of valid and reliable instruments and methodologies. Gordon and Sweeney[6] have proposed three models for identifying and validating diagnostic labels: clinical, retrospective, and nurse validation. Fehring[5] suggested that diagnostic labels should be examined for their diagnostic content validity, clinical diagnostic validity, and etiologic correlation ratings prior to being accepted for use in clinical practice. Validating nursing diagnoses is not an easy process, because validity must be viewed as a question of degree rather than an all-or-none characteristic. Research provides support for but not proof of validity, thus making validation a neverending process that requires the use of multiple techniques rather than a single methodologic strategy. The purpose of this article is to discuss four sequential phases for examining the content and construct validity of nursing diagnoses: the Delphi technique, magnitude estimation scaling, patient observation, and nursing intervention studies.[7] These proposed methods represent a potentially useful strategy in addressing nursing's responsibility for the diagnosis and treatment of human responses to actual or potential health problems.

PHASE I: DELPHI TECHNIQUE

Determining the validity of a nursing diagnosis requires several sequential steps that build on each other in a logical fashion. The first phase consists of determining the sampling adequacy of defining characteristics and operational definitions for a nursing diagnosis. One methodologic approach for determining content validity of defining characteristics and operational definitions for a nursing diagnosis is the modified Delphi technique. The Delphi technique is a special type of survey method for soliciting and obtaining group consensus on an important subject. This technique was first developed as a forecasting tool at the Rand Corporation in Santa Monica, California, in the 1940s and has since been applied to economics, industry, research, business, and education for problem solving, planning, forecasting, and evaluation.[8] Key characteristics of this technique are (1) anonymity of panel members, (2) anonymity of responses, (3) numerous iterations, (4) group consensus, (5) controlled feedback of responses to panel members, and (6) statistical analysis of data.[8,9]

The procedure consists of asking a panel of experts in a specific content area to evaluate a set of items over a series of written questionnaires. The number of rounds in a Delphi varies and may be influenced by the level of consensus desired by the researcher, concreteness or abstractness of the items, serendipitous data, and number of items or length of the questionnaires. The average number of rounds required to reach an acceptable level of consensus on a majority of items, however, is usually three. Sometimes a summarization of round 3 as well as written comments is considered to be round 4.[9]

Using the modified Delphi technique to examine the content validity of the defining characteristics for nursing diagnoses is a critical step in the validation process, because the development and findings of this

phase will influence the examination of construct validity in sequential phases. Errors as well as successes in the development of defining characteristics and their operational definitions for a nursing diagnosis become evident in future phases. Therefore a panel of content experts is required to validate the defining characteristics and the operational definitions for a diagnostic label. In this phase the panel of content experts validates the association of proposed characteristics with the diagnostic label and the accuracy of operational definitions for each defining characteristic. It is important that the panel of experts includes researchers in other disciplines who are engaged in inquiry related to the phenomenon under study (eg, sleep, pain).

The format of the Delphi may vary based on the type of diagnostic label and number of defining characteristics and operational definitions. However, certain requisites must be met when using the Delphi technique. In the first round content experts judge the appropriateness and clarity of defining characteristics and operational definitions for a nursing diagnosis. The proposed defining characteristics and operational definitions are derived from literature pertinent to a specific nursing diagnosis. Space should be provided for the content experts to make comments regarding why a defining characteristic or operational definition is inappropriate and unclear and to make suggestions for revising defining characteristics or operational definitions. Also space should be allotted for the content experts to identify additional defining characteristics and operational definitions for the proposed nursing diagnosis. Based on experience with the process, the authors recommend that the maximum number of defining characteristics and operational definitions initially identified by the investigator should be no more than 45. Return rates of questionnaires through several rounds may be low and is even poorer when the questionnaire is lengthy.[10] In addition, the ability of the content experts to remember previous rounds is diminished with numerous defining characteristics. Finally, a diagnostic label with a very large number of defining characteristics may be best explained by more than one nursing diagnosis.

The development of clear and concise defining characteristics and operational definitions is important in employing the Delphi technique. The actual format of the Delphi may vary by using either a four-point Likert-type scale or yes/no responses when judging the appropriateness and clarity of the characteristics. Use of a four-point scale reduces the likelihood that responses of the content experts will regress toward the mean and may yield more meaningful data if the nursing diagnosis is initially defined by a small number of defining characteristics. On the other hand, use of a yes/no format may be more appropriate for a larger number of defining characteristics. Regardless of the level of agreement achieved, it is recommended that no defining characteristics and operational definitions be deleted from the list prior to the completion of two rounds. Two rounds are usually required to develop clear and concise defining characteristics and operational definitions and to eliminate colloquial terms peculiar to one geographic location.

The percentage of agreement in retaining a defining characteristic or operational definition is set by the investigator and will be influenced by sample size and level of abstraction of the nursing diagnosis, defining characteristics, and operational definitions. Although an acceptable percentage of agree-

ment may range from 70% to 100%, it is suggested that a relatively high percentage of agreement be set initially, such as 85% or 90%.[9,11] This level of agreement will serve to increase feedback from multiple exposures so that clearer and more appropriate defining characteristics and operational definitions may be developed for a specific nursing diagnosis. It is recommended that the minimal acceptable level of agreement for the final round of the Delphi should be 70%.

PHASE II: MAGNITUDE ESTIMATION SCALING

The second sequential phase in determining the validity of a nursing diagnosis consists of examining the construct validity of the proposed label. Magnitude estimation is a methodologic strategy that shows promise as a scaling technique in examining the construct validity of nursing diagnoses. By adapting psychophysical methodology to social stimuli and subjective responses, magnitude estimation scaling (MES) yields ratio level measures by assigning numbers to stimuli that are proportional to the magnitude of an individual's subjective responses.[12,13]

According to Hinshaw,[14] MES requires the identification of a set of clearly and concisely stated stimuli. Examples of stimuli include a list of objects, events, people, items on an instrument, or defining characteristics for a nursing diagnosis. General psychometric rules for constructing items for scales also apply when identifying defining characteristics for use with MES. Not only must each defining characteristic be clear and concise, but it must also contain only one major cue. For example, the defining characteristic "inconsistent verbal and motor responses" contains two cues related to verbal and motor responses, while "inconsistent verbal responses" and "inconsistent motor responses" contain only one major cue in each defining characteristic. The set of defining characteristics must also vary on the major concept dimension to be scaled to evoke varied responses. Some concept dimensions that may be used to validate nursing diagnosis are importance of the defining characteristic, frequency of occurrence of the defining characteristic, or competence of the nurse in identifying the defining characteristic. For example, if the concept dimension of importance is to be scaled using a set of defining characteristics, then the investigator needs to include defining characteristics that are of little importance, mild importance, moderate importance, and great importance in making the nursing diagnosis.

A second requirement of the procedure is the development of operational definitions for the concept dimensions used to scale the defining characteristics. The operational definitions should contain only one broad, conceptual dimension to stimulate subjective responses. Hinshaw[14] recognized the difficulty in simultaneously developing precise attitudinal variables that measure only one conceptual dimension and are broad in scope. Following is an operational definition for the concept dimension of importance that meets this criterion: When assessing a patient with (specify nursing diagnosis), some signs and symptoms are more important than others in making this nursing diagnosis. For example, more weight or significance is assigned to critical signs and symptoms.[7,15]

Operational definitions are needed also for the defining characteristics when using MES to validate nursing diagnoses. Delineating operational definitions for defining characteristics not only facilitates the scaling of defining characteristics on concept dimensions, but also provides a consistent standard

for observing patients in the clinical setting. Operational definitions should range widely on a continuum of severity to facilitate identifying patients who have different degrees of the defining characteristics.

Another requirement of this procedure is an adequate number of subjects who possess appropriate knowledge of expertise to make proportional judgments. Knowledge or expertise for MES is influenced by cultural and educational conditioning.[14] Potential respondents for validating a nursing diagnosis through the use of MES are nurses who have in-depth knowledge of defining characteristics that confirm a nursing diagnosis (eg, nurses working in a neuroscience intensive care unit would judge the importance of defining characteristics for altered level of consciousness). Hamblin and Smith[16] suggested that 20 to 30 subjects are adequate for MES.

The fourth and final requirement of MES is the training of subjects to make proportional judgments when presented with a set of stimuli such as defining characteristics. Lodge[12] recommends a procedure in which subjects match lines to numbers and then numbers to lines. These calibration tasks are essential, since some subjects would make ordinal judgments without training. The assignment of numbers to line-length stimuli, or numeric estimation, is the most commonly used calibration task, because it is a stable exponent across a wide range of stimuli. Drawing lines to a set of number stimuli, or line production, also has a fairly stable exponent.

In the training sessions, subjects practice assigning numbers to lines that are proportional to a standard or reference line. The reference line should be divisible by many numbers, such as a 60-mm reference line. Hamblin[17] suggests an easy training procedure to verify the subjects' ability to make proportional judgments. Below is an example of numeric estimation training.

> Let's practice making proportional judgments with numbers. Here is a booklet that contains a set of line lengths. The first line in the booklet is your reference and has been assigned the number 60. Please note that some of the lines in the booklet are longer than the first line and some are shorter. Your task is to judge how much shorter or longer each line is compared to the reference line by assigning each line a number that is in proportion to 60. For example, if a line is twice as long as the reference line, assign the number 120. On the other hand, if a line appears to be one half as long as the reference line, assign the number 30. If a line appears to be the same length as the reference line, assign the number 60.

Subjects also practice drawing lines in relation to numbers that are in proportion to a reference number such as 60.[12] Below is an example of line production training.

> Let's practice making proportional judgments with lines. Here is a booklet containing a set of numbers. The first number in the booklet is 60 and corresponds to a line length that will serve as your standard. I would like you to draw lines that are in proportion to the standard. For example, the number 120 is twice as large as the number 60. Therefore you would draw a line twice as long as the standard. You would draw a line one half as long as the standard for the number 30.

After verification that the subjects are able to make proportional judgments, the subjects score each defining characteristic for a particular nursing diagnosis according to one or more concept dimensions, such as importance or frequency of occurrence. For example, from a set of randomly ordered defining characteristics, subjects are asked to select one characteristic that is judged to occur with average frequency in confirming the nursing diagnosis under investigation (eg, altered sleep pattern). The characteristic that the subject judges as occurring with average frequency in confirming altered sleep pattern is assigned a number (eg, 60) and serves as a reference or standard. Then subjects assign numbers to the other characteristics in proportion to the characteristic that is the standard. For example, if the subject judges one characteristic as occur-

ring twice as frequently as the standard in confirming the nursing diagnosis, it is assigned the number 120. If the subject judges another characteristic as occurring one fourth as frequently as the standard in confirming the diagnosis, it is assigned the number 15. It is possible and even likely that subjects will select different standards from the list of defining characteristics. Following is an example of instruction given to subjects when scaling defining characteristics on the concept dimension of frequency.

Here is a set of defining characteristics for the diagnosis of (specify nursing diagnosis). I would like you to judge how frequently these defining characteristics occur when making the nursing diagnosis. Take a few minutes to read the defining characteristics and their operational definitions. Now select a sign or symptom that you judge to occur with "average" frequency in making the nursing diagnosis. This characteristic will serve as a standard and will be assigned the number 60. Now score how frequently the other defining characteristics occur in relation to the defining characteristic you have selected as a standard. For example, if a defining characteristic occurs twice as frequently as the standard in making the nursing diagnosis, assign it a score of 120. If a defining characteristic occurs just as frequently as the standard, assign it a score of 60. If the defining characteristic occurs one third as frequently as the standard, assign it a score of 20.

Analysis of magnitude estimation scores yields frequencies, ranges, medians, and geometric means. Raw scores are logarithmically transformed so that linear rather than curvilinear statistical procedures may be performed. Data analysis is based on the principle of averaging, whereby random error is removed, thus providing a more accurate measure of the variables.[13] Pearson's product-moment correlations are calculated to determine relationships between concept dimensions, such as the relationship between importance and frequency of occurrence of defining characteristics. Percentiles or quartile ranges are used in distinguishing between critical and less important characteristics. It is suggested that the upper quartile be used to isolate critical defining characteristics of a nursing diagnosis.

Reliability coefficients for MES are high. Hinshaw and Murdaugh[18] reported test-retest reliability coefficients of $r = .96$ to .98 with 90 nurse subjects who used MES to judge 13 nurse activities (stimuli) according to their complexity. In estimating construct validity, correlations ranged from $r = .94$ to .99 when nurses judged the complexity of nursing activities on four modalities: MES, line production, hand dynamometer, and sound pressure.

PHASE III: PATIENT OBSERVATION

The third phase is further examination of the construct validity of a nursing diagnosis through patient observation. Gordon and Sweeney[6] proposed that validation of nursing diagnosis emerges from occurrence of the diagnostic label in clinical practice. Several methodologic concerns must be addressed in the clinical phase: choosing a standard to guide selection and to compare groups of subjects, constructing an instrument, controlling for extraneous variables, determining sample size, conducting a pilot study, determining intrarater or interrater reliability, selecting appropriate statistical tests, and identifying critical defining characteristics based on the findings of phases II and III.

The selection of a standard for identifying and comparing subjects requires careful deliberation. Ideally the standard should be unobtrusive, objective, and practical.[19] An excellent example of a standard with all of these features is the Glasgow Coma Scale (GCS) developed by Teasdale and Jennett.[20] The GCS has been widely used in clinical practice to assess level of consciousness and could be employed to identify subjects with and without altered level of consciousness as well as for assigning subjects to comparison

groups. While other standards may be more elusive, several sources may be helpful in this endeavor: nursing care plans, expert judges, standardized instruments, tests, or patient verification.[5,6,21]

Choosing a standard is an essential step in the clinical phase to discriminate between subjects who do and do not have a nursing diagnosis. A standard also serves as a constant frame of reference so that the same criteria are used to select subjects. Choosing a standard to select subjects must be based on the literature. For example, the literature frequently suggests that level of consciousness consists of three areas: best eye opening, best motor response, and best verbal response. Ideally a standard used to select subjects with and without an altered level of consciousness would address these three components.

Following the selection of a standard, an instrument must be constructed to observe patients with and without the nursing diagnosis. The instrument should provide space to record demographic variables (eg, age, sex, medical diagnosis) as well as group assignment information that would allow the investigator to differentiate among groups when analyzing data (eg, GCS scores). Operational definitions should be placed on the instrument in close proximity to the list of defining characteristics for a nursing diagnosis to facilitate consistency of data collection. The actual format used to record data will vary based on whether the variables are discrete or continuous. For example, some defining characteristics are usually recorded as either absent or present rather than on a continuum. Therefore it would be most appropriate to record these types of defining characteristics as either absent or present on the data collection sheet. In contrast, defining characteristics such as "discomfort" may be measured best by use of a Likert-type scale, because discomfort is frequently reported by degree of severity. Providing a column on the instrument that is designated as "nonapplicable" is appropriate for those situations in which defining characteristics cannot be observed. Space should be allotted on the instrument to record significant comments such as the presence of an endotracheal tube.

Controlling for extraneous variables is important in patient observation and is based on the review of literature. An extraneous variable obscures the accurate measurement of defining characteristics. For example, neuroscience literature supports that fluid and electrolyte imbalances, acid-base imbalances, hypoglycemia, uremic or hepatic encephalopathies, and exogenous agents such as drugs or poisons influence level of consciousness.[22,23] Controlling for these variables may eliminate many potential subjects. Controlling for these variables, however, does not mean that extraneous variables must be completely absent in the patient. For example, patients in intensive care units frequently have electrolyte imbalances. The investigator must decide what degree of imbalance in fluids and electrolytes will affect the observation of certain defining characteristics for an altered level of consciousness such as disorientation or muscle weakness. This expert judgment is derived from clinical practice and a review of the literature. For example, the investigator may accept abnormal potassium levels of 3.2 to 5.4 mEq/L (normal is 3.5 to 5.2 mEq/L) in a patient in an intensive care unit, because disorientation and muscle weakness do not usually occur when potassium levels are in that range.

Sample size is influenced by the number of defining characteristics for a nursing diag-

nosis, the number of groups, the type of sampling procedure, and the type of statistical test used to analyze data. Because it may not be possible to observe for the absence or presence of some defining characteristics (eg, verbal response when an endotracheal tube is present), the investigator may need to compensate by increasing the sample size.

A pilot study is essential prior to initiating the research to judge the appropriateness of the instrument designed by the investigator. Determining intrarater or interrater reliability is a part of the pilot and should continue at various points during patient observation to ensure consistency of data collection. The number of data collectors will depend on several factors in patient observation. Factors influencing patient observation and reliability are the level of abstraction of the defining characteristics and whether the absence or presence of defining characteristics may change in a short period of time. For example, because level of consciousness may change in a patient in a short period of time, it would be more appropriate to determine interrater reliability with two or more individuals observing the same patient with an altered level of consciousness. Using more than one individual to collect data will also decrease investigator bias.

Selecting the appropriate statistical test will depend on the type of research question, level of measurement of the data, sample size, and number of groups. For example, a Fisher's exact probability test (if data are nominal, $n < 30$, and expected value in each cell is < 5) or X^2 test of independence (if data are nominal, n 30, and expected value in each cell is 5) may be used to determine whether the frequency of occurrence of defining characteristics for a nursing diagnosis was significantly different when comparing patients with and without a particular nursing diagnosis. When using a Likert-type scale, a *t* test may be an appropriate statistical test to compare groups if basic assumptions of the test are met (eg, normal distribution, homogeneity).

The data from phases II and III must be analyzed concurrently to determine critical defining characteristics. Critical defining characteristics are those that achieve geometric means in the upper quartile range and occur with significantly greater frequency in patients with the nursing diagnosis. In contrast, defining characteristics that are not appropriate for the nursing diagnosis would achieve geometric means in the lower quartile range and would not occur with significantly greater frequency in patients with the nursing diagnosis.

Nursing diagnoses must have clinical referents, that is, defining characteristics must appear with significantly greater frequency in patient populations with the nursing diagnosis when compared with populations without the nursing diagnosis. It is essential that the proposed critical defining characteristics match patient behavior in the clinical setting prior to studying the nursing diagnosis for its usefulness in directing nursing interventions and predicting patient outcomes.

PHASE IV: NURSING INTERVENTION STUDIES

Validation of diagnostic labels is not complete until nursing diagnoses have been studied to determine their usefulness in directing nursing interventions and in predicting patient outcomes. According to Friedman and colleagues,[24] the most powerful experimental technique for assessing the

effectiveness of an intervention is a properly planned and executed clinical trial.

In a clinical trial intervention techniques are applied in a standard fashion, and response variables are measured to answer the study question. While a full description of the clinical trial is beyond the scope of this article, some important decision points will be addressed.

Primary and secondary questions must be formulated and will most often be posed as hypotheses. The primary question might ask whether the response variable is altered by the intervention, while secondary questions might address the relationship of certain patient characteristics to the response variable. For example, music therapy might be applied to patients with a diagnosis of Alteration in Comfort: Pain with the response variable of interest being analgesic use. Secondary questions of interest might relate age or gender to analgesic use.

The intervention is designed to maximize the potential benefit that might be demonstrated while minimizing any untoward effects. Because nursing interventions frequently require interactive techniques, attention must be given to the standard application of the intervention and to the control of extraneous variables.

Response variables, or patient outcomes, must be carefully selected to represent appropriate criteria of efficacy and to be clinically meaningful. The response variable may be discrete or continuous. For example, analgesic use may change from use to nonuse or from one level of use to another.

Eligibility of subjects must be determined in advance using well-defined criteria. Inclusion criteria control the homogeneity of the groups and influence the confidence placed in any demonstrated differences between the control and experimental groups.

A number of study designs are available, but it is important to recognize the advantages of random assignment of subjects to groups in detecting effects of the intervention. Potential problems of bias during the conduct of the trial may be handled by blinding techniques, although these may be difficult with selected diagnoses.

It is of utmost importance that baseline data be acquired before the intervention is applied to evaluate group comparability. The variables should be selected with care and may include physical, psychosocial, or demographic characteristics. The quality of the data is a key concern throughout the study and requires that sufficient attention be devoted to avoid problems.

While there are many issues to be considered relative to data analysis, follow-up, interpretation of findings, and collaborative trials, excellent texts are available to assist the nurse researcher in investigating the effect of a nursing intervention in patients with an identified nursing diagnosis.[24-26] No study is without flaw but attention to study questions, intervention, outcome, inclusion and exclusion criteria, design, and other features of the clinical trial will maximize the information that can be gained and will contribute to the advancement of clinical nursing knowledge.

• • •

A methodologic approach consisting of four phases has been proposed for validating nursing diagnoses. The sequential phases include use of the Delphi technique, magnitude estimation scaling, patient observation, and testing nursing interventions. It is not recommended that all of these phases be completed in one research study but rather

that the phases be conducted sequentially. The proposed methodologic approach is a strategy for researching the diagnosis and treatment of human responses to actual or potential health problems—the foundation for nursing practice.

REFERENCES

1. Aspinall MJ. Nursing diagnosis—the weak link. *Nurs Outlook*. 1976;24(7):433–436.
2. Putzier DJ, Padrick KP. Nursing diagnosis: a component of nursing process and decision making. *Top Clin Nurs*. 1984;5(4):21–29.
3. Jacobs M, Huether SE. Nursing science: the theory-practice linkage. *ANS*. 1978;1(1):63–73.
4. Kelly MA. *Nursing Diagnosis Service Book: Guidelines for Clinical Application*. Norwalk, Conn: Appleton-Century-Crofts; 1985.
5. Fehring RJ. Validating diagnostic labels: standardized methodology. In: Hurley ME, ed. *Classification of Nursing Diagnosis: Proceedings of the Sixth Conference*. St Louis, Mo: Mosby; 1986.
6. Gordon M, Sweeney MA. Methodological problems and issues in identifying and standardizing nursing diagnosis. *ANS*. 1979;2(1):1–16.
7. Kinney M, Guzzetta C. Identifying critical defining characteristics of nursing diagnoses using magnitude estimation scaling. *Res Nurs Health*. 1989;12(6):373–380.
8. Notter L. *Essentials of Nursing Research*. 3rd ed. New York, NY: Springer; 1983.
9. Lindeman C. Delphi survey of priorities in clinical nursing research. *Nurs Res*. 1975;24:434–441.
10. Seaman C. *Research Methods: Principles, Practice, and Theory for Nursing*. 3rd ed. Norwalk, Conn: Appleton & Lange; 1987.
11. Polit D, Hungler B. *Nursing Research: Principles and Methods*. 2nd ed. Philadelphia, Pa: Lippincott; 1983.
12. Lodge M. *Magnitude Scaling: Quantitative Measurement of Opinions*. Beverly Hills, Calif: Sage; 1981.
13. Sennott-Miller L, Murdaugh C, Hinshaw A. Magnitude estimation: issues and practical applications. *West J Nurs Res*. 1988;10:414–424.
14. Hinshaw A. Role attitudes: a measurement problem. In: Hardy M, Conway M, eds. *Role Theory: Perspectives for Health Professionals*. Norwalk, Conn: Appleton-Century-Crofts; 1978.
15. Grant J. *Altered Levels of Consciousness: Validity of a Nursing Diagnosis*. Birmingham, Ala: University of Alabama at Birmingham; 1989. Unpublished doctoral dissertation.
16. Hamblin R, Smith C. Values, status, and professors. *Sociometr*. 1966;29:183–196.
17. Hamblin R. Social attitudes: magnitude measurement and theory. In: Blalock H Jr, ed. *Measurement in the Social Sciences*. Chicago, Ill: Aldine; 1974.
18. Hinshaw A, Murdaugh C. *Validity of Ratio Measurement of Nursing Concepts: A Cross Modality Study*. Washington, DC: US Dept of Health and Human Services Division of Nursing; grant No. 00971. 1985.
19. Rudd P. In search of the gold standard for compliance measurement. *Arch Intern Med*. 1979;139:627–628.
20. Teasdale G, Jennett B. Assessment of coma and impaired consciousness. *Lancet*. 1974;2:81–83.
21. Fitzmaurice J. Nurses' use of cues in the clinical judgement of activity intolerance. In: McLane M, ed. *Classification of Nursing Diagnoses: Proceedings of the Seventh Conference*. St Louis, Mo: Mosby; 1987.
22. Plum F, Posner J. *Diagnosis of Stupor and Coma*. 3rd ed. Philadelphia, Pa: FA Davis; 1980.
23. Ricci M. Neurological examination and assessment of altered states of consciousness. In: Nikas D, ed. *The Critically Ill Neurosurgical Patient*. New York, NY: Churchill Livingstone; 1982.
24. Friedman LM, Furberg CD, DeMets DL. *Fundamentals of Clinical Trials*. Littleton, Mass: PSG Publishing; 1985.
25. Spilker B. *Guide to Clinical Studies and Developing Protocols*. New York, NY: Raven Press; 1984.
26. Spilker B. *Guide to Planning and Managing Multiple Clinical Studies*. New York, NY: Raven Press; 1987.