Methodological Problems and Issues in Identifying and Standardizing Nursing Diagnoses

Marjory Gordon, R.N., Ph.D., F.A.A.N.
Professor of Medical-Surgical Nursing
School of Nursing

Mary Anne Sweeney, R.N., Ph.D.
Associate Professor, School of Nursing
Boston College
Chestnut Hill, Massachusetts

CENTRAL to the development of a clinical science is the description of the phenomena of concern. In nursing, identifying health problems diagnosed and treated by nurses is the primary step in organizing and building a clinical science. When health problems are clearly delineated, studies can focus upon the natural history, therapeutic management and outcomes of each problem. Historically, most of the emphasis in clinical nursing research has been on task-oriented functions such as decubitus treatment, pain alleviation or patient teaching. In most disciplines, however, isolating the phenomena of concern through descriptive research precedes experimental studies designed to alter or explain the phenomena. As Kritek suggests, prescriptions and prescriptive theory have been generated without the prior stages of development.[1]

In 1961 Abdellah et al. published a list of 21 nursing problems.[2] The list consisted of therapeutic goals such as "To maintain

2

good hygiene and physical comfort." This promoted the idea of classifying nursing problems or patients' nursing needs but did not isolate the clinical phenomena or health problems that could provide the structure for systematic development of a clinical science. A significant step in delineating the elementary structure and scope of a clinical science in nursing, albeit not an intentional focus, is the current, national effort to identify and classify nursing diagnoses. Three National Conferences on Classification of Nursing Diagnoses have yielded approximately 37 diagnostic category areas.[3-5] Nurses from all geographic areas of the United States and Canada viewed these as problems they treated in their clinical practice.

This work of the National Conference group represents an approach to the step of specifying and describing the phenomena in the field of nursing and can provide the potential elements of structure for a clinical science. Further work is needed to expand this listing and validate the diagnoses. This type of clinical research is associated with a number of conceptual and methodological problems and issues.

CONCEPTUAL ISSUES

Two major conceptual issues are encountered in attempts to design research in the area of nursing diagnosis. One concerns how the term can be operationalized with the current diversity which exists in the conceptualization of nursing. The second issue concerns the deductive versus inductive approach to identifying diagnostic categories.

Operational Definition of Nursing Diagnosis

Defining nursing diagnosis is one of the major conceptual issues confronting the researcher. A definition requires structural, conceptual and competency components. The structural component refers to the observable characteristics, the conceptual component delineates the focus or type of health problem included as a nursing diagnosis and the competency component describes the type of people who make nursing diagnoses. Gordon includes structural and competency referents in a definition of nursing diagnosis.[6(p1298)] It is viewed as a concise term representing a cluster of signs and symptoms and describing an actual or potential health problem or state-of-the-patient which nurses, by virtue of their education and experience, are licensed and able to treat. The structural component is further specified. A nursing diagnosis contains the problem, etiology, signs and symptoms. The competency component of this definition relies heavily on education, individual competencies and nurse practice acts, none of which provide the uniformity required. The necessary structural elements of a diagnosis are clear but the domain of problems addressed is not.

Other definitions refer to the domain of health problems in abstract terms. Soares[7(p275)] views diagnosis from a conflict-in-needs framework, Yura and Walsh[8(p114-115)] view it as deprivations or alterations in meeting human needs and Little and Carnevali[9(p46)] see it as a response to a stressor or potential stressor in the health area. The most concise definition

combining structural and conceptual elements is that proposed by a group of nursing theorists.[5] They have defined nursing diagnosis as "a concise phrase or term summarizing a cluster of empirical indicators representing patterns of unitary man." Even these few examples demonstrate the lack of consensus on whether the conceptual focus of nursing diagnosis should be needs,[8] stress responses,[9] patterns of unitary man,[10] self-care deficits[11] or adaptations.[12] Although there is little unanimity regarding a nursing model, most nurses agree that the medically oriented model of biological systems and pathological states is not adequate for nursing.

Using the definition, "nursing diagnosis is a product of assessment," the participants of the three national conferences have identified diagnoses in the areas listed in Table 1. Although lacking an explicit conceptual focus for their work, the participants have identified and proposed clinical definitions for diagnoses that most would agree are within the nursing domain.

The diagnoses now accepted require validation and additional diagnoses need to be identified. If subjects are asked to identify or validate nursing diagnoses, the concept must be clearly defined by the investigator. The solution to this problem

3

TABLE 1
Areas of Diagnostic Nomenclature Currently Accepted

Anxiety	Nutritional alteration
Bowel elimination, alteration in	Parenting, alterations in
Cardiac output, alteration in	Respiratory dysfunction
Comfort, alterations in	Role disturbance
Consciousness, altered levels of	Self-care activities, alterations in
Coping patterns, maladaptive (individual)	Self-concept, alterations in
Coping, ineffective family	Sensory-perceptual alteration
Digestion, impairment of	Sexuality, alteration in patterns of
Family process, inadequate	Skin integrity, impairment of
Fear	Sleep-rest activity, dysrhythm
Fluid volume, alteration in	Spirituality, matters of
Grieving	Social isolation
Home maintenance management, impaired	Suffocation, potential for
Injury, potential for	Thought processes, impaired
Knowledge, lack of (specify)	Tissue perfusion, chronic abnormal
Manipulation	Trauma, potential for
Mobility, impairment of	Urinary elimination, impairment of
Noncompliance (specify)	Verbal communication, impairment of

Sources: Gebbie, K. M. and Lavin, M.A., eds. *Proceedings of the First National Conference: Classification of Nursing Diagnoses* (St. Louis: The C.V. Mosby Co. 1975).

Gebbie, K. M., ed. *Summary of the Second National Conference on Classification of Nursing Diagnoses* (St. Louis: Clearinghouse for Nursing Diagnoses 1976).

Gebbie, K. M., ed. *Proceedings of the Third National Conference on Classification of Nursing Diagnoses* (St. Louis: April 1978), to be published in 1980.

4 or issue is a model that is relevant to nursing in its theoretical abstraction and that is sufficiently concrete to be used in clinical research.

In medicine, for example, the model is system or organ dysfunction and a diagnosis describes a disease, syndrome or specific set of observations. The compilation of 999 medical diagnoses listed in the International Classification of Diseases presents a clear example of the clinical entities which have been catalogued by physicians to date.[13] The nursing profession is in the formative stages of a similar task.

The Inductive versus the Deductive Method

A second conceptual issue in identifying diagnostic nomenclature is the choice between the inductive and deductive methods. The inductive method involves the description of clients' health problems by nurses, either during their practice as participant-observers or from their memory of client problems encountered in practice. The deductive method does not require a clinical setting for identifying diagnoses. A conceptual model—e.g., patterns of unitary man—is made explicit and diagnoses are logically deduced by nurses familiar with the model.

The participants at the National Conferences utilized the inductive method.[3-5] During these conferences, participants described and labeled the patient problems encountered in their practice. The majority argued against acceptance of one conceptual model as the underpinning for their work. This practice-oriented approach has led successfully to the identification of problems or problem areas, most of which can be validated as being taught in nursing curricula. Other groups of practitioners agree that they are within the scope of nursing practice.

As the list of nursing diagnoses increased, the National Conference group perceived a need for a conceptual model to organize, classify and direct further identification. Currently, a group of nursing theorists are developing a conceptual framework, based on unitary man, to guide the National Conference group's work.[5] If a concrete, clinically usable model can be agreed upon, further diagnoses can be generated deductively. For example, if a set of categories representing patterns of unitary man can be specified, common patterns requiring nursing intervention can be identified and labeled in diagnostic terms. Validating that the diagnoses occur clinically can follow. This would provide a description of the clinical phenomena and the structural elements of a clinical science.

Many would argue that some implicit conception of nursing guided the National Conference participants' inductive identification of health problems amenable to nursing therapy. Similarly, the inductive approach, when used in clinical identification of nursing diagnoses, also requires operationalizing some concept of nursing. Otherwise, no guide is available to determine which clinical data are processed for purposes of nursing diagnoses and treatment and which data are collected for purposes of referral to physicians or other health care providers.

Most nursing theorists involved in the

With either the inductive or deductive approach, diagnostic labeling should be concise, clear and clinically useful.

identification of nursing diagnoses have argued for a deductive approach from a conceptual model. Using a model such as self-care deficits, diagnoses would be generated, labeled and tested for clinical usefulness. The problem with the deductive approach is deciding which nursing model would generate diagnoses that would be acceptable to all nurses.

With either the inductive or deductive approach, diagnostic labeling should be concise, clear and clinically useful. For example, the phrase "alterations in self-concept" is too abstract and vague for clinical usefulness. It may be a broad classification for organizing specific diagnoses, but it is too encompassing a term for determining specific interventions. However, broad diagnostic categories, when employed clinically, may lead to further discrimination of specific categories when relevant data are accumulated. In medicine, for example, "dropsy" was a diagnosis until further study led to the identification and discrimination of congestive heart failure and a number of renal diseases.

Two types of studies will be needed to eventually develop a classification system of clinical diagnoses. Studies focusing on identification and labeling are the first type required. Although diagnostic labels could be derived solely from the nursing literature, this would require the assumption

that it contains adequate descriptions of all health problems currently treated by nurses. Secondly, research investigations are required to validate the nomenclature as it develops. This type of project requires large-scale testing to determine the clinical occurrence of diagnostic entities and consensual validation of diagnostic labels. Following this research, national or international standardization can occur. Diagnoses can be continually added to the initial listing.

The considerable undertaking described requires the consolidation of research findings from a great number of diverse studies focused on varied aspects and types of nursing diagnoses. Valuable time and energy can be conserved by meticulous planning as well as through the utilization of current work in the field. The remaining sections of this paper have been constructed to provide guidelines in two important aspects of this clinical research process: the selection of a research model and an explication of some of the key issues in these types of investigations.

MODELS FOR IDENTIFICATION AND VALIDATION OF NURSING DIAGNOSES

The literature on nursing diagnosis has not emphasized research models, although three basic models for identifying and validating diagnostic nomenclature can be proposed using principles of descriptive research. These prototypes are classified as the retrospective identification model, the clinical model and the nurse-validation model.

6   *The Retrospective Identification Model*

The retrospective identification model utilizes the accumulated experiences of nurses. Subjects are directed to verbalize descriptions of health problems they have treated in the past. These verbal data are assumed to be based on their memory of clinical encounters with patients and on their accrued nursing knowledge.

Sample selection, using this method, must be given careful consideration. Decisions regarding subjects' years of nursing experience (particularly direct patient care experiences), clinical specialty knowledge, familiarity or expertise in diagnosis and geographic spread must be carefully weighed. If a particular health problem area is the focus of study—e.g., roles-relationships—consideration should be given to a cross-specialty sample of nurses. Many health problems such as mild anxiety which are routinely treated by nurses are common to many of the nursing clinical specialties.

Another consideration is stratification and randomization of the sample. This may be less desirable in diagnostic identification studies than in studies for purposes of standardization where representative, consensual validity is of concern.

The procedure for a retrospective identification study is similar to that used by the participants at the National Conferences.[3(p45)] Approached from a research design perspective, a large, geographically representative sample of educators and practicing nurses votes on a set of diagnostic labels shortly after they are generated from small group work sessions. Initially, small groups of nurses are given a health care or health problem area such as nutrition or one of the subsets of a conceptual model. They are then directed to identify the specific problems diagnosed and treated within the broad framework. Guidelines for constructing a diagnostic category which includes the problem-label, etiology, signs and symptoms (defining characteristics) would be necessary, as would guidelines for arriving at large group consensus. The Delphi technique may be feasible for this type of methodology when geographic limitations preclude a face-to-face encounter.

Consensual validity can be tested on the diagnoses generated by this method if a second sample of nurses with representative characteristics is used. When provided with the cluster of signs and symptoms (defining characteristics) of the diagnosis, do they apply the same or similar labels? To decrease variability in responses, guidelines on the process of labeling should be provided. This will decrease some of the semantic variability in the use of words.

The reliability and validity attained in a retrospective model depend to a great extent on the investigator's sample selection and training methodology. As described above, the subjects should be scrutinized carefully on the basis of such factors as experience, expertise and clinical background. The amount and type of training received in the definition and

The reliability and validity attained in a retrospective model depend to a great extent on the investigator's sample selection and training methodology.

elements of nursing diagnosis may influence the project outcomes.

The advantage of the retrospective model is the feasibility of obtaining a large, representative sample of nurses. Disadvantages are inherent due to the possible heterogeneity of the sample members who are identifying and validating the diagnoses. It may be that adaptations of the procedures in this model and the initiation of a study incorporating the Delphi technique would provide a method for national standardization of diagnoses already identified.

An adaptation of this model was used in the Third National Conference.[5] Diagnoses generated by individuals and regional work groups were reviewed and voted upon in relation to specified criteria for a diagnostic category. The 100 to 200 nurses who participated in the voting process at each conference represented geographical regions of the United States and a variety of nursing positions and specialties.

The Clinical Model

The clinical model uses direct observations of patient behaviors as the source of diagnostic labels. Such studies are carried out in clinical settings with an emphasis on verbal data from the nurse-data collector or written data from patient records. Use of this model of course necessitates consideration of patient rights and institutional access.

Sample selection requires the same consideration as the retrospective model using nurse-subjects. The investigator, however, is more dependent on the ability of the nurses to collect and process clinical data as well as on their facility to communicate accurate, clear descriptions of clinical phenomena. Unless simultaneously or subsequently carried out in settings across the country, geographic representation, stratification and randomization are sacrificed.

The time and effort employed in designing even a local study of this nature argue for employing training and reliability procedures. The heterogeneity of the nurses who act as participant-observers must be reduced by training sessions followed by measures of diagnostic reliability. The key aspects of studies of this type are the nurse's assessment and diagnostic skills.

The required tools include: assessment guidelines based on a conceptual model of the client, guidelines for diagnostic categories, protocols for entering and discharging patients from the study and a format for data collection. If the data collectors are not project staff, their responsibilities to their employers must be weighed. Change in established agency protocols for patient assessment and the nonresearch routine of nurse-patient activities on the unit have to be considered as potential problem areas prior to initiation of the study.

The selection of a health problem area or age group from which diagnoses can be generated is a major consideration in the study design. It is advisable to select an area which is not too restrictive. For example, the focus of a study could be a population of patients with newly acquired mobility or activity-tolerance problems. Irrespective of the focus, instead of assessing one limited area of human behavior, a complete initial patient assessment should

8

be incorporated into the protocol. This will assure the availability of a full range of diagnostic information. The time and cost of this type of study should be associated with a later payoff in the number and quality of diagnoses identified. Castles' data from an intensive care unit suggest that preliminary identification studies should not be carried out in settings where patients have high medical needs and physiological instability.[14] Areas where the greater percentage of health problems are tended to by nurses may generate more data—e.g., convalescent, community or long-term care settings.

Holmstrom and Burgess employed the clinical model for diagnostic category identification using a population of patients undergoing a crisis experience.[15] Participant-observer nurses collected biopsychosocial data during postrape counseling sessions. The nurses were trained in assessment and counseling and used a flexible interview guide. The investigators analyzed the assessment data and labeled the clusters, generating three nursing diagnoses. The large sample provided an indication of the incidence of each diagnostic category in the sample population.

The Nurse-Validation Model

Following identification of diagnostic categories but prior to standardization, the cluster of signs and symptoms that define the diagnosis should be validated. The term(s) or label(s) are merely an abbreviated, arbitrary description. Of major concern is whether the cluster of defining characteristics to which the label refers, actually occurs as an entity in clinical

situations. Thus validation of a diagnosis involves determining if the preidentified defining characteristics occur as a cluster in a sufficient number of cases. Whether nurses agree on the appropriateness of labels for the clinically validated clusters is of secondary concern.

Clinical validation is especially important when labels and defining characteristics are identified using the retrospective model, which employs recall of clinical

---

Clinical validation is especially important when labels and defining characteristics are identified using the retrospective model, which employs recall of clinical practice experiences.

---

practice experiences. The nurse-validation model can be applied to the clinical testing of identified diagnoses. As with previous models, the nurse-subjects who are participant-observers in the study need to be trained and their reliability as diagnosticians established. In addition, they need to be familiar with the diagnostic labels to be tested.

The procedure involves tabulating which defining characteristics listed for a diagnosis are present when the diagnosis is made in clinical practice. Additional signs and symptoms observed by the nurse-subjects should also be recorded. For example, with a sample of 100 patients diagnosed as having a particular nursing diagnosis, the investigator could determine the actual frequency of each identified defining characteristic (sign or symp-

tom). This would serve as a measure of the coexistence or validity of the cluster of patient characteristics defining the diagnosis. High frequency characteristics would be the criterial cluster for defining the diagnostic label.

Kim developed a design for use within a clinical setting. In this design two nurses do assessments on the same patient sample and make diagnoses independently using a list of identified diagnoses.[16] Another methodology entails using a panel of specialists in both nursing and the diagnostic process to review the data collected by trained diagnosticians. Consensual validity could also be obtained by this latter method. Campbell used undergraduate students' care plans to collect data and a clinical panel to review the data on the basis of set criteria for a diagnosis.[17] Reliability and validity measures were not reported.

The nurse-validation model could also be employed in settings where problem-oriented or computerized record systems are utilized and in which signs and symptoms, etiologies and nursing diagnoses are recorded. Data can be culled from nurse-subjects' recordings to eliminate several steps of the data collection procedure. Although it is not common to find patient-related recordings which are sufficient for research purposes, this state of affairs will change when nurses become more cognizant of clinical research, the need for data retrieval and the value of their clinical observations and judgments.

M. Gordon, R.N., Ph.D., M. A. Sweeney, R.N., Ph.D. and K. McKeehan, R.N., M.S. used this methodology in a study of the frequency of discharge diagnosis in a group of gynecological and obstetrical patients referred for continuing care (unpublished data, 1979). No interrater reliability studies were done on the judgments of the nurses making the diagnoses, yet the hospital had training sessions in the use of a list of diagnoses for the nurse diagnosticians. All discharge diagnoses were reviewed in a team conference by a continuing care coordinator who was expert in nursing diagnoses. Thus some degree of consensual validity in labeling and diagnostic reliability was provided.

THE USE OF RATERS

Clinical models for identification and validation which utilize nurses' judgments about nursing diagnoses have been mentioned. Investigators, such as Castles, have utilized highly specialized nursing practice sites, such as critical care units, and have asked the nurses to diagnose patients' problems.[14] Subjects were not provided with a list of suggested categories for identifying client problems or diagnostic training, thus it was necessary to rely upon individual judgment and wording preference of each nurse. Kim designed a study which utilized baccalaureate students to identify client problems.[18] Each participant was provided with a list of possible diagnoses, assembled by the National Conference group and was directed to use the list as well as any additional self-generated labels. The common thread running through these types of investigations is the use of an observer or rater to collect and analyze clinical signs and symptoms and to label client prob-

10

lems requiring nursing care. Two types of issues need to be considered in depth: the readiness of the rater to make a reliable judgment and the manner in which the judgment process in enacted.

The use of raters or judges as part of the research process is a long established procedure. Investigators who are planning studies which involve raters must ensure that the raters' judgments adhere to acceptable levels of validity, sensitivity, reliability and clinical-reality orientation. This requires training raters and establishing their level of diagnostic reliability, as well as reporting these data to permit pooling of studies in this vital area. Investigators, utilizing raters to identify nursing diagnoses, should carefully consider: (1) the nursing expertise of the raters, (2) specialized training in using nursing diagnosis and (3) methods of selecting diagnostic labels such as a prepared list versus the use of completely open-ended responses.

The experience of the authors in the implementation of a five-month pilot project for training diagnosticians (raters) may be of value to other investigators. The project was an educational exercise to provide information on some of the anticipated problems in training and utilizing raters prior to the initiation of a more formalized research investigation. Using the project as an example will highlight the multifaceted nature of the issues which are involved in this aspect of preparing for research.

The Educational Project

One of the authors presented a two-hour class on Research Issues in Nursing Diag-

noses to senior level baccalaureate nursing students during a required research course. A group of 24 students expressed an interest in learning more about the topic of nursing diagnoses and sought further information about participating in a research project which was being planned. Each student expressing such an interest was invited to join a study group. The group had approximately ten meetings over a four and one-half month semester. Seven students completed the diagnostician training project as an extracurricular activity.

The first meeting of the trainees was held approximately one and a half months after the classroom introduction to nursing diagnosis. The project training sessions each consisted of three parts: (1) a discussion of the theoretical aspects of nursing diagnoses, (2) rating sessions in which the students were asked to use simulated clinical data in formulating nursing diagnoses and (3) discussions of the students' formulated diagnoses to provide immediate feedback and clarification as needed.

The discussion of theoretical aspects consisted of clarification of points raised by the authors and students alike. Each rater was provided with a set of articles which were considered key works in the field, as well as an extensive bibliography of additional references. A limited number of audio tapes was used within the training sessions, although tapes of important lectures on the topic were placed in the library for individual use. All the student volunteers were motivated to expand their knowledge base in this area and were accustomed to reading reference material in conjunction with baccalaureate course assignments. They reported that the read-

ing and discussion of the published articles played an important part in the development of a clear understanding of the diagnostic process as well as in the formulation of a nursing framework for making judgments.

Recognition and identification exercises were used in the rating sessions. A recognition exercise consisted of nursing diagnoses mixed with a list of symptoms, therapeutic needs, treatments and medical diagnoses. Thus practice in discriminating between negative and positive instances of the concept of nursing diagnosis was provided. Small vignettes constructed from the signs and symptoms of a diagnosis were used in identification exercises. The clinical vignettes utilized in the initial stages of the project for the selection and application of nursing diagnoses were fairly simple and brief. During the course of the rating sessions the degree of difficulty increased with the use of lengthy and involved actual patient case examples. The authors varied the process of selecting diagnostic labels. In some cases, possible diagnostic labels were set forth in a multiple choice fashion and some were left completely open-ended. In other cases or vignettes the raters were instructed to choose a label from the list of diagnoses accepted by the National Conference group.[3-5]

The materials used in rating sessions were being used simultaneously in workshops. This provided some consensual validity for the diagnostic labels used in cases and vignettes. Conflicting answers were resolved by the authors who were familiar with nursing diagnosis and with the cognitive aspects of the diagnostic process. Discussion of individual diag-

noses, feedback and clarification followed the case and vignette exercises. Feedback and the rationale for correct and incorrect diagnoses were initially given by the authors. As training progressed, the students spontaneously began to offer this information to each other.

Although it is not possible to make definitive statements about the effect of specific aspects of the process, some of the trends that became apparent during training are worth noting. Agreement among raters, agreement between raters and experts (the authors) and the method used to formulate diagnoses are the most essential aspects of the training process to be considered.

A method for assessing agreement on cases or vignettes was piloted. Agreement was defined as concurrence on the major focus of a diagnostic label. For example, although the terms differed, "lowered self-esteem" and "decreased self-esteem" were considered synonymous. It was found that interrater agreement alone was inadequate as a measure of diagnostic reliability.

The initial agreement among raters was quite high during the first two months of training (79% and 75%), as may be seen in Table 2. The level of agreement was sharply reduced by the third month of the project (40% and 50%) and slowly increased until the final high level of 88% agreement among raters was reached at the conclusion of the training sessions. The percent of agreement between the raters and experts showed a somewhat different pattern. The original agreement level of 79% (attained on two vignettes with multiple choice options) dropped off sharply while the raters attained a higher level of agreement with each other (55% with

11

TABLE 2
Summarization of Diagnostic Rating Sessions

Date	Number and Type of Cases	Raters Participating Labeling Exercise	% Agreement Among Raters	% Agreement Between Raters and Experts	Type of Labeling Selections Utilized
2/8/78	2 brief	1-7	79	79	Multiple choice
3/20/78	4 brief	6,4,5,1,7	75	55	Open ended, no list
4/19/78	1 extensive	1-7	40	45	Open ended, no list
4/26/78	1 extensive	1-4	50	75	Open ended, no list
5/3/78	9 brief	3,1,5,4,6	82	82	National Conference list
5/8/78	2 brief	1-7	63	63	Open ended, no list
5/8/78	4 brief	6,4,5,1	88	88	National Conference list

experts versus 75% with other raters) during the second rating session. The raters then started showing higher agreement with the experts than with each other (45% versus 40% and 75% versus 50%) until similar levels of agreement between both groups was reached in the fourth month of the project.

The causes of the different trends appear to be related to two factors: (1) an understanding of the concept of nursing diagnosis and (2) experience in its use. The initial trainees' ratings yielded diagnoses which tended to be medically oriented. Labels such as diabetes mellitus were chosen rather than ones describing nursing problems. This caused the lowered rate of agreement between the raters and the experts during most of the first phase. As the trainees became more familiar with the process of isolating health problems in terms of nursing diagnoses, agreement with experts and with each other increased noticeably. This two-way agreement is an essential aspect of rater training.

Although the trainees exhibited a fairly high level of interrater reliability in the initial phase, the meaning of this agreement level was recognized only when it was placed in its proper perspective—that is, compared to the judgments of the experts. Trainees were adept at picking out relevant clinical data but their statement of the problem was within a medical or symptomatic focus. Thus, ratings need to be correlated with a relevant outside criterion.

A plausible explanation for the trend in both sets of agreement ratings is the new knowledge and skill gained through the experience. The opportunity to read a number of carefully selected articles and to discuss various points about the articles and their relationship to clinical practice undoubtedly had beneficial effects on the evolving process of making diagnostic judgments. The time devoted to the discussion of the nursing diagnoses selected by each rater for the test cases was undoubtedly a strong influence on the

evolving agreement scores. The opportunity to share views, to receive immediate feedback on inappropriate label selection, to identify individual idiosyncracies (using too broad a perspective) as well as group problems (struggling with nomenclature on cases where diagnoses were made in an open-ended fashion) fostered the processing of clinical data in a more identical way. Raters need training programs and discussion sessions to ensure that they are approaching the task with a common framework and are making judgments from a common basis. They also need to validate their judgments with an outside criterion, such as experts.

The use of varying exercises in labeling highlighted some interesting points regarding both level of difficulty and semantic variability. The vignettes with multiple choice responses were rated in the shortest span of time. This method also provided for clear-cut evaluations of differences between correct and incorrect responses. However, it can produce deceptively high agreement since the number of options is limited and the level of difficulty lowered. Open-ended vignettes in which the trainee has to generate diagnostic labels also have advantages and disadvantages. Vignettes using the open-ended response can produce lowered agreement because of semantic differences in the nomenclature used and variation in the level of abstraction of the diagnostic responses (broad diagnostic categories as opposed to specific diagnoses). The exact diagnostic terms used can be expected to vary with this method, although the problem area identified is similar. This exercise may be most productive for groups of more

sophisticated raters or those with training in the structure of diagnostic labels.

The third method of designating diagnostic labels entailed utilization of a list of nursing diagnoses identified by the National Conference group.[3-5] This type of approach encouraged the trainees to consider the total range of diagnostic categories produced to date and seemed to help them narrow options rather than generate them by chance. It was comprehensive, less time consuming than working with no list at all and provided labels which could be discussed with ease. Interestingly, the National Conference nomenclature was easily grasped by these senior baccalaureate students. They appeared to know the meaning of the terms, requiring only a label for their mental concept. The use of a list of nursing diagnoses produced the highest level of agreement but it also was used following training sessions in recognition and identification which introduced terminology.

The pilot training project served to identify some of the variables influencing rater reliability in this area: (1) structure of the exercises, (2) provision of a nomenclature listing and (3) the inherent difficulty of the diagnoses and diagnostic clusters (signs and symptoms combination) used in the exercises. It can be expected that these factors will also vary with the pretraining diagnostic expertise of the potential rater.

THE NEED FOR NURSING DIAGNOSIS

Research is critically needed to identify and validate health problems, thereby

14

organizing the body of clinical science in nursing. Nursing diagnoses can form the basis for writing outcomes in quality assurance programs. They can assist in the measurement of the "cost effectiveness" of the nursing care rendered by professionals and may be utilized in varied managerial tasks such as computing staff assignments.

In the field of education, nursing diagnoses can provide the structure or focus of the content of clinical science.

In the field of education, nursing diagnoses can provide the structure or focus of the content of the clinical science. Lastly, diagnoses within the scope of nursing practice can be the focus of clinical nursing research. For example, predictive studies can be conducted, interventions can be studied and compared, outcomes can be correlated with presenting problems and prognoses can be empirically identified.

Lack of conceptualization of the research model or utilization of questionable research procedures can result in tenuous research results. This may be one of the reasons that no large-scale studies have been funded in nursing diagnoses despite the widespread interest. Development and testing of research designs for identification and standardization of nursing diagnoses is critical. The American Nurses' Association Standards of Practice, which require diagnoses, and the Nurse Practice Acts[19] which mandate nursing diagnoses, make it imperative that research in this area move forward.

REFERENCES

1. Kritek, P. "The Generation and Classification of Nursing Diagnoses: Toward a Theory of Nursing." *Image* 10:2 (June 1978) p. 33-40.
2. Abdellah, F. et al. *Patient-Centered Approaches to Nursing* (New York: MacMillan 1961).
3. Gebbie, K. M. and Lavin, M. A., eds. *Proceedings of the First National Conference: Classification of Nursing Diagnoses* (St. Louis: The C. V. Mosby Co. 1975).
4. Gebbie, K. M., ed. *Summary of the Second National Conference on Classification of Nursing Diagnoses* (St. Louis: Clearinghouse for Nursing Diagnoses 1976).
5. Gebbie, K. M., ed. *Proceedings of the Third National Conference on Classification of Nursing Diagnoses* (St. Louis: April 1978) to be published in 1980.
6. Gordon, M. "Nursing Diagnoses and the Diagnostic Process." *Am J Nurs* 76:8 (August 1976) p. 1298-1300.
7. Soares, C. A. "Nursing and Medical Diagnoses: A Comparison of Variant and Essential Features" in Chaska, N., ed. *The Nursing Profession* (New York: McGraw-Hill Book Co. 1978).
8. Yura, H. and Walsh, M. B. *Nursing Process* 3rd ed. (New York: Appleton-Century-Crofts 1978).

9. Little, D. and Carnevali, D. "Diagnostic Statement: Problem Defined" in Walter, J. B., Pardee, G. P. and Molbo, D. M. *Dynamics of Problem-Oriented Approaches: Patient Care and Documentation* (Philadelphia: J. B. Lippincott Co. 1976).

10. Rogers, M. *Introduction to the Theoretical Basis for Nursing* (New York: F. A. Davis 1979).

11. Orem, D. E. *Nursing: Concepts of Practice* (New York: McGraw-Hill Book Co. 1971).

12. Roy, Sr. C. *Introduction to Nursing: An Adaptation Model* (Englewood Cliffs, N.J.: Prentice-Hall 1976).

13. World Health Organization. *Manual of the International Statistical Classification of Diseases, Injuries and Causes of Death* Vol. I, II (Geneva: WHO 1977).

14. Castles, M. E. "Study of Nursing Diagnoses Generated in an Intensive Care Setting" in Gebbie, K. M., ed. *Proceedings of the Third National Conference on Classification of Nursing Diagnoses* (St. Louis: April 1978) to be published in 1980.

15. Holmstrom, L. and Burgess, A. W. "Development of Diagnostic Categories: Sexual Traumas." *Am J Nurs* 75:8 (August 1975) p. 1288-1297.

16. Personal communication with Kim, M. J., Nursing Faculty, Illinois University School of Nursing, June 1977.

17. Campbell, C. *Nursing Diagnosis and Nursing Interventions in Nursing Practice* (New York: John Wiley & Sons 1978).

18. Kim, M. J. "Utilization of Nursing Diagnoses by Baccalaureate Nursing Students" in Gebbie, K. M., ed. *Proceedings of the Third National Conference on Classification of Nursing Diagnoses* (St. Louis: April 1978) to be published in 1980.

19. Bullough, B. "Law and the Expanding Nursing Role" in Chaska, N., ed. *Nursing Profession: Views through the Mist* (New York: McGraw-Hill Book Co. 1978) p. 315-317.