Strategies For Building Nursing Databases
For Effectiveness Research

Judy G. Ozbolt, PhD, RN, FAAN

Recognized as an urgent priority in investigations of health services delivery, effectiveness research applies epidemiological methods to large databases to determine the relative costs and efficacy of alternative modes of treatment and to explore the effects of covariates and intervening variables on relationships among problems, interventions, outcomes, and costs. For reasons of quality, economics, and sound management, nursing, like other health professions, must undertake effectiveness research. Unfortunately, however, although nursing leaders have long recognized the importance of constructing databases from nursing's clinical information for purposes of research and nursing management, and although they have for many years urged researchers to study relationships among structure, process, and outcome in nursing, nurse scholars still have not developed the methodological tools and resources to conduct effectiveness research.

The prime requirement for effectiveness research is utterly lacking in nursing: valid, reliable, useful data from multiple cases and multiple sites structured in databases that allow easy, preferably automated entry from clinical records and ready access for whatever analyses the investigators may choose to perform. The absence of specific nursing data from national and regional databases is not, however, the result of discrimination against nursing. It is rather the result of the profession's failure to agree upon and offer a set of clearly defined, valid, reli-

able, and standardized data elements for inclusion. To achieve such a data set is a formidable undertaking, but a number of useful efforts are already underway.

Of central importance, the elements of the Nursing Minimum Data Set have been agreed upon. Furthermore, most of them are already included in the Uniform Hospital Discharge Data Set. The missing items are the unique identifiers of patient and principal nurse, the addition of which would appear to be primarily a matter of policy and politics, and the four nursing care items, Nursing Diagnosis, Nursing Intervention, Nursing Outcome, and Intensity of Care. It is these four latter items that present difficulties with regard to validity, reliability, utility, and standardization.

Considerable work has been done to develop data elements in each of the four categories, but validity, reliability, and utility are in question and comprehensiveness and standardization are lacking. Strategies are proposed for developing items in each category and testing their validity, reliability, and utility for effectiveness research.

The key point is that, however urgently nursing may require inclusion of its data in large databases for effectiveness research, the development and selection of the data items must be as soundly scientific as the subsequent research using the data. Rushing to consensus without adequate development and testing would risk invali-

dating the effectiveness research based on the data. The proposed strategy would intersperse multi-site research projects to develop and test data elements with invitational conferences of nurse scholars engaged in this endeavor to share progress and make recommendations for further research and development. Consensus on items for inclusion in databases for effectiveness research could thus be based on research and scientific debate. Although this strategy might take longer than forcing items willy-nilly into the databases, the seriousness of the consequences cries out for careful science. Invalid or unreliable data could obscure real patterns and lead to false conclusions. And effectiveness research using the databases will inform policy decisions that profoundly affect the practice of nursing and the nation's health.

Introduction

In 1966, Avedis Donabedian called for assessing the quality of patient care along three dimensions: structure, process, and outcome. The 1970's saw a flurry of nursing research activity directed toward quality of care, including efforts at the University of Wisconsin—Madison (Zimmer et al., 1974), and The University of Michigan (Horn & Swain, 1977) to establish patient outcome measures that would reflect the quality of nursing care. During that decade also, nursing leaders gave high priority to the need for research to document relationships among structure, process, and outcomes in nursing. Nurse researchers did not, however, take up this challenge, perhaps in part because of the relative immaturity of nursing research programs generally. In the 1980's, the costs and intensity of care rose relentlessly, more nurses and more highly skilled nurses were required in all patient care settings, and resource expenditure became a critical issue. Public demand to justify the high cost of care peaked as the decade ended, and the dawn of the 1990's brought a new urgency to the need for research to demonstrate the effectiveness and the cost-effectiveness of nursing care.

This research must show not only *that* nursing makes a difference, but *how* nursing affects patient outcomes, and how much it costs.

The National Center for Nursing Research is seeking to establish a program of patient outcomes research. To that end, participants have gathered at this conference to review the state of the science and to make recommendations for research directions. Presenters have discussed the importance of effectiveness research in nursing. They have described approaches to defining, identifying, and measuring interventions and outcomes and examined the effects of factors other than nursing care on patient outcomes. And they have discussed existing databases, with the opportunities and the shortcomings they present for research on nursing effectiveness.

All these deliberations bring us to a startling conclusion. Twenty-five years after Donabedian showed us the way to examine the issues of quality of care, we still are not ready to carry out effectiveness research in nursing. We do not have the methodological tools and resources. To respond, finally, to the challenge to demonstrate links between structure, process, and outcomes in nursing—and make no mistake, our response will profoundly affect whether nursing thrives or withers in the coming decades—we must answer three questions:

o   What is required to carry out effectiveness research in nursing?

o   What do we have now to build on, and where are there gaps in our methodologies?

o   How can we get from where we are to where we need to be?

Requirements for Effectiveness Research in Nursing

Effectiveness research uses epidemiological methods to examine large databases. Investigators may target specific problems and look for their precursors or associated factors, the various interventions used to treat the problems, the results achieved, and the costs. They may also look for covariates and intervening variables in the relationships among problems, interventions, outcomes, and costs. In addition, they may explore such issues as whether different subgroups of the population are diagnosed and treated differently, with different costs and different outcomes.

The most obvious requirement for such research is the databases themselves. The databases must contain large numbers of cases, preferably from multiple sites, and their data elements must be standardized. Furthermore, the data elements must be *valid* representations of the underlying constructs: patient problems, interventions, outcomes, and resource expenditures. In addition, the data must be *reliable*, and the standard of reliability should be higher than that usually accepted in clinical practice. In practice, one inaccurate data point is unlikely to have a drastic effect on the patient's welfare, because clinicians will consider the total picture the patient presents in making decisions about care. In research, by contrast, if a data item is regularly collected inaccurately, this unreliability will make it difficult to detect real patterns and may lead to false conclusions. Finally, these valid, reliable data from multiple cases and multiple sites must be structured in databases that allow easy, preferably automated entry from clinical records and ready access for whatever analyses the investigators may choose to perform.

Current Resources for Effectiveness Research

Essential Data Elements

To date, nursing data have been excluded from large databases such as the Uniform Hospital Discharge Data Set. This is not due to any discrimination against nursing, but rather to the profession's failure to agree upon and offer a set of clearly defined, valid, reliable, and standardized data elements for inclusion. To achieve such a data set is a formidable undertaking, but a number of useful efforts in that direction are already underway.

For more than 30 years Harriet Werley has been urging nurses to attend to the issues of collecting, managing, and processing clinical data in ways that will facilitate communication, decision making, and research. In 1985, in conjunction with Dean Norma M. Lang of the University of Wisconsin—Milwaukee School of Nursing, she convened a working conference to develop an initial Nursing Minimum Data Set (Werley & Lang, 1988). Invited participants prepared papers exploring a number of considerations in the development and implementation of a nursing minimum data set, and working groups deliberated the specification and definition of elements. Refined by a Post-conference Task Force, the work of the Conference Group resulted in a list of 16 elements in three categories. There are four Nursing Care Items, five Patient or Client Demographic Items, and seven Service Items. Apart from the Nursing Care Items (Nursing Diagnosis, Nursing Intervention, Nursing Outcome, and Intensity of Nursing Care) and two Service Items (Unique Health Record Number of Patient or Client and Unique Number of Principal Registered Nurse Provider), all the items are already included in the Uniform Hospital Discharge Data Set (UHDDS) (Health Information Policy Council, 1985). Incorporation of the Nursing Minimum

Data Set into the standard regional and national databases will provide precisely the data needed for effectiveness research in nursing.

So why do we not act immediately to achieve the inclusion of the Nursing Minimum Data Set in these databases? To get the missing Service Items included in the UHDDS would seem to be primarily a matter of policy and politics, with consideration for the privacy and security of information. The Nursing Care Items, however, are still awaiting the establishment of their validity, reliability, and utility as a prerequisite to developing data standards.

Nursing Care Items

Nursing diagnoses. The first Nursing Care Item, Nursing Diagnoses, is the most highly developed, but even this item requires further work before it will be ready for standardization. Werley and Lang (1988, p. 405), reporting on the work of the Conference Group and the Postconference Task Force, commented that

> This designation [of nursing diagnoses] is to include all diagnoses identified from admission through discharge. These nursing diagnoses are to be drawn from diagnoses accepted through the systematic review process of the North American Nursing Diagnosis Association (Kim, McFarland, & McLane, 1984), or from other recognized diagnostic systems.

The North American Nursing Diagnosis Association is still developing what is intended to be a comprehensive set of diagnoses. Some areas of nursing practice do not yet have established diagnoses, and the existing diagnoses vary in clinical specificity. Furthermore, there has been little research to support the defining characteristics of accepted diagnoses.

Other individuals and groups who have developed sets of diagnoses specific to their populations, settings, or purposes include Campbell (1984), Joel (1984), Ver Steeg (1988), Hastings and Muir-Nash (1989) and the Executive Committee of the Division of Psychiatric and Mental Health Nursing Practice of the American Nurses' Association (Carroll-Johnson, 1989). For more than a decade, the Omaha Visiting Nurses Association has been developing and refining a set of diagnoses and, more recently, interventions and outcomes, related to community health nursing practice (Simmons, 1980; Visiting Nurses Association of Omaha, 1986; Martin, Scheet, Crews, & Simmons, 1986). These various endeavors are essential to developing the clinical language for the issues nurses address in their practice, but they have not yet resulted in a set of diagnoses sufficiently comprehensive to be representative of the full scope of nursing practice.

In addition to language and definitional problems, measurement problems also remain to be resolved. While the correlations between the diagnoses and their putative defining characteristics are still being explored, the more basic issue of interobserver reliability in detecting the presence or intensity of the defining characteristics awaits empirical demonstration. Existing assessment tools with tested validity and reliability (e.g., Horn & Swain, 1977; Fillebaum & Smythe, 1981; and Waltz & Strickland, 1988) do not pretend to include all the issues that might be the object of nursing care. Thus, despite extensive efforts by multiple investigators, nursing is far from possessing a valid and reliable assessment guide for establishing the presence, absence, or degree of a comprehensive set of nursing diagnoses.

Nursing interventions. Similarly, there is no universally accepted language or set of definitions for nursing interventions, although in the past decade several investigators have made intensive efforts in this area. In 1984, Campbell proposed a way of categorizing some 2500 nursing interventions for use in nursing information systems. Finding that number unwieldy and perceiving the need for an orderly classification system, Bulechek and McCloskey (1985, 1987) and their associates at the University of Iowa are developing a taxonomy of nursing interventions that will be tested for content validity. Their methods include sampling intervention statements and related activities from nursing textbooks, care planning books, and similar sources and inductively developing a classification system and a taxonomy of interventions. Concurrently and independently, Grobe (1989) and colleagues at the University of Texas at Austin have been collecting nursing intervention statements from nurses, analyzing them to produce a nursing intervention lexicon, and building a taxonomy. And, as mentioned above, the Omaha Visiting Nurses Association includes standardized language for nursing interventions in its information system (Martin, Scheet, Crews, & Simmons, 1986). The degree to which these and other efforts may complement or conflict with each other and collectively produce a comprehensive taxonomy of nursing interventions has not yet been examined. Thus, in developing standardized language for nursing interventions, as for nursing diagnoses, the profession has come a long way but still is far from the goal.

Nursing outcomes. In the Nursing Minimum Data Set, Nursing Outcome is defined as "Resolution status of the nursing diagnosis(es)" (Werley & Lang, 1988, p. 403). Conceptually, then, the language and measurement issues are much the same as those for diagnoses. The nurse would be assessing the same phenomenon at a different point in time and, ideally, a healthier point on the health-illness continuum. But if the phenomena are taken to be the same as the nursing diagnoses, measuring them entails all the same issues of validity and reliability, plus some specific to the condition of being outcomes of nursing care. How, for example, does one define a clinically important change in status? And how does one deal with the issue of multiple causality of outcomes? Measures of nursing outcomes developed in the 1970's (e.g., Zimmer and colleagues, 1974; Horn & Swain, 1977) should be as reliable today as they were then, but their validity is in question. The items developed by Zimmer and colleagues (1974) were specific to the medical treatments and the anticipated course and length of hospitalization of the period; some if not most would be outdated today. The Horn and Swain instrument (1974) identified patient outcomes that nurses validated as achievable with good care in a time when the average hospital patient had a less acute illness and a longer stay than today. With a few notable exceptions, such as the work of the Omaha Visiting Nurse Association (Martin, Scheet, Crews, & Simmons, 1986), more recent efforts to develop measures of patient outcomes have tended to focus on specific diseases or conditions and so lack the comprehensiveness needed for effectiveness research on all the clients of a type of health care agency (see, e.g., Waltz & Strickland, 1988). Developing language and measurement standards for outcomes of nursing care thus lags behind diagnoses and interventions. A great deal of work must be done before nursing outcomes can become valid and reliable clinical data abstractable for effectiveness research.

Intensity of nursing care. Many automated systems have been developed to classify patients according to the amount of nursing care required. Because these data are available, nurses tend to rely on them for some indication of relatively greater or lesser cost of care. They have certain disadvantages, however. One is lack of sensitivity. Although these classification systems can pro-

vide a rough estimate of overall expenditure of nursing resources on a patient or on a unit, they are not designed to keep track of the resource expenditure related to particular interventions, nor are most of them linked to the actual care plan or documentation. Thus, if nurses were comparing two different methods of caring for patients with the same problem, those patients might be rated the same in "acuity;" if that were the only measure of resource consumption, any cost savings from one of the methods of care would be invisible. Another problem is the lack of standardization in the systems themselves, so that data are not comparable across institutions and may not be comparable across units. The problem is not limited to systems from different vendors; different agencies may implement the same system in different ways and so produce non-comparable data. Although the beginning efforts to validate the Nursing Minimum Data set include attempts to standardize assessment tools and classification schemes, there is still a dearth of research into methods of determining the cost of nursing care, research needed to guide the development of data standards for effectiveness research.

Strategies for Progress

The current situation, then, is that nursing has a pressing need for large databases amenable to effectiveness research in nursing, using epidemiological methods. Furthermore, in the Nursing Minimum Data Set, nurses have proposed the *kinds* of data elements needed for such databases, and few if any nurses have opposed the proposition or suggested alternatives. The difficulties are in defining, measuring, and standardizing the specific variables to be included within each category of data elements. *Which* diagnoses, interventions, outcomes, and estimates of intensity should be included? *How* should they be measured? And how recorded?

Given the economic and political urgency of using effectiveness research to demonstrate the value of nursing and the practical need to find the most effective and least expensive ways to deliver care, it is very tempting to rush to consensus. Nurses might be inclined to review existing approaches to defining and measuring the data elements and, by some process of professional group dynamics, to select one of each to define and measure diagnoses, interventions, outcomes, and intensity. This would be a grave mistake. Lacking any scientific basis for the decision, nurses would risk adopting data definitions and standards that would obscure the real impact of nursing care, as detectable through the methods of effectiveness research. At the same time, potentially superior approaches might be prematurely abandoned. More fundamentally, if we are to develop the science of effectiveness research in nursing, we must also be scientific in constructing its foundations, the data elements on which all future research will depend.

What scientific methods might we use to arrive at definitions, measures, and standards for the data elements needed for effectiveness research?

Nursing Diagnoses

Validity. The North American Nursing Diagnosis Association (NANDA) has a consensual process for adopting diagnoses that serves as a check on validity. Others may wish to consider a variety of approaches. For example, nurses practicing within a particular clinical specialty might consider that the current set of NANDA diagnostic labels did not adequately cover their practice. Investigators could use a combination of interviews, observation, and debriefing to determine the issues these nurses address in their practice and the terms they use to designate them. By methods similar to those Grobe (1989) is using for interventions, researchers could analyze the

statements to develop a lexicon and then a taxonomy of nursing diagnoses for that area of specialty practice. A useful criterion for the inclusion of a diagnosis would be that its recognition results in specific action; otherwise, it would simply name a patient condition that nurses do *not* address, and ought not to be included in a list of nursing diagnoses. Extending the research to different institutions and regions would provide a rich sample of local colloquialisms and help investigators to determine which usages of terminology are so widespread that they should become the standard, which regional variations should also be considered acceptable, and which idiosyncratic usages should be discouraged. Carefully done,this approach could lead to terminology that captures the language of everyday practice and standardizes it for universal recognition. Extending such research across a broad range of specialty areas would allow the set of diagnoses to approach comprehensiveness and should elucidate which, if any, diagnoses are common to nursing practice across specialty lines. Constructing a taxonomy of these terms would illustrate their relationships—for example, by showing which terms were a subset of which others—and would clarify issues of level of abstraction and clinical specificity.

Identifying the issues nurses address in their practice and developing language to express them, important as it is, is only a part of validity. Investigators must also determine the validity of the defining characteristics by which nurses recognize the presence or degree of a diagnosis. By interviewing nurses just after they have made a diagnosis, investigators might determine which cues led to the diagnosis, and so determine the frequency of occurrence of specific cues with diagnoses and the strength of relationships. Researchers must bear in mind, however, that this method will uncover only those cues that come to conscious thought and can be articulated, and that the recognition of a cue might depend as much on the nurse's level of expertise as on its actual presence in a situation. Still, conscientious research should produce, for each diagnosis, a list of defining characteristics with some indication of the strength of their association with the diagnosis.

Reliability. The unequal ability of nurses to recognize a cue or a diagnosis raises the issue of reliability. As investigators identify diagnoses and related defining characteristics, they must examine and refine the ways of assessing and measuring them. They must also determine interobserver reliability for each of the data items. The method and standard used by Horn and Swain (1977) offer sufficient rigor for research use of the data that meet the standard while being demonstrably attainable for a wide variety of clinical data. In this method, two researchers collect data on the same patient simultaneously and independently for subsequent analysis of reliability on each item. The criteria of reliability are a .80 index of agreement, statistical significance at the .05 level, and confidence limits that exclude the .60 criterion at the lower bound. Items that do not meet these criteria should be revised, retested, and if necessary discarded. Implementation of these criteria of reliability would in itself contribute to improved patient care, as nurses could more accurately recognize diagnoses. Furthermore, effectiveness researchers could be confident that the diagnoses recorded in the database did, in fact, occur in the patient.

Nursing Interventions

Validity. At least two federally supported research projects are currently underway to develop taxonomies of nursing interventions (Bulechek & McCloskey, 1985, 1987; Grobe, 1989). Both of these include plans to determine validity. Their methods are different, and their products are likely to be different. When they are complete, investigators should study them

via the methods of content analysis. First, they should examine them for similarities and differences. If they should be found to have identified the same interventions in some areas, this would provide validation that both practicing nurses and the literature agreed that certain interventions were characteristic of nursing practice. If they have identified different interventions in some areas, do they complement or conflict with each other? If they seem to have used different language to express approximately the same idea, can they be reconciled? Separately and together, how comprehensive are they of the universe of nursing interventions? From such careful analysis of the results of these projects, illuminated by professional discussion and debate, can arise conclusions about the validity and completeness of the taxonomies and recommendations for further research.

Reliability. With interventions, the reliability question is whether different nurses will apply the same label to the same nursing action. Interrater reliability studies should proceed with similar criteria for acceptance of a label as those recommended for acceptance of diagnoses. Problems might arise if there were different labels for very similar interventions, or if the terminology were so obscure that the nurse did not immediately recognize the label as applying to the action. The way a particular information system presented the data might also make it relatively easier or harder for the nurse to locate and choose the correct label. Investigators studying interrater agreement would need to consider all these factors in accounting for instances of disagreement.

Nursing Outcomes

Validity. If nursing outcomes are defined as the "resolution status of nursing diagnosis(es)" (Werley & Lang, 1988, p. 403), then the language to express them must parallel the language of the nursing diagnoses. In-

vestigators working to develop and test nursing diagnoses might apply the same methods to generating and testing statements of outcomes of nursing care. This implies that finding language to express patient problems is not sufficient. Rather, nurse researchers must identify the underlying issue that is the focus of nursing concern and that may manifest as a problem (as we usually think of diagnoses), or as a strength (as we hope the outcomes will be), or as some point on a continuum between the most negative and the most positive possible conditions. Because outcomes are linked to diagnoses, their validity is linked to the validity of the diagnoses. That is, if the diagnosis represents an issue that nurses address in caring for the patient, then the outcome ought logically to be influenced by the nursing care given. And indeed, in the Nursing Minimum Data Set, nursing outcome is also defined as, "An aspect of patient or client health status that is influenced by nursing intervention and recorded at specific times for an episode or encounter of care" (Werley & Lang, 1988, p. 407). Content validity of both diagnoses and outcomes could be established by having nurses review items and indicate whether the issue is indeed one they address in practice, and whether they would be willing to have the quality of their care evaluated by the patient's outcome status on the item. Although many factors, including the patient's own recuperative powers, contribute to the outcomes actually achieved, patients who receive excellent nursing care ought, in the aggregate, to achieve better outcomes than do patients who receive care that is mediocre or worse. Otherwise, the issue or problem is not sufficiently influenced by nursing care to qualify as a nursing diagnosis, and the outcome is not the result of nursing. Such items should not be part of a nursing data set used for effectiveness research. That set should be reserved for items representing issues that nurses act upon and whose outcomes they influence.

Reliability. As the validity issues related to outcomes parallel those related to diagnoses, so do the reliability issues. If diagnosis and outcome represent different points along the same continuum, the measurement concerns are quite similar. Investigators should be aware of the possibility of such difficulties as non-uniform sensitivity in scales and measuring instruments; that is, the tool that adequately measures a condition in its problematic or "diagnosis" stage may or may not be sensitive enough to assess the outcome status; and vice versa. Interobserver reliability of outcome items can be determined by the same methods and criteria as those recommended for diagnoses.

Intensity of Care

Validity. Automated systems to calculate patient acuity or intensity of nursing care have generally been extensively tested for the validity of their estimates of overall nursing time consumed by a patient or a set of patients, as a guide to staffing. Whether any or all of them are valid indicators of resource consumption for effectiveness research is an open question. There is at least one situation where they clearly are not. These systems base their estimate of time required for patient care on some version of "standard practice." The assumptions on which the classification system is based remain the same even when the care varies. Particularly for effectiveness research, when nurses may wish to compare the costs and outcomes of two or more different interventions or models of care, it is critical to be able to measure costs (or resource consumption) on the basis of what is actually occurring, rather than relying on some standard estimate. With automated systems that include documentation of care delivered and the person providing it, it should be possible to calculate resource consumption more directly and sensitively from the clinical record. Nurses should recognize the important advantages that such di-

rect calculations would offer (including not having to buy, maintain, or collect data for a separate "acuity" system) and insist that vendors and institutional officials make this kind of system available.

Reliability. The lack of standardization of existing acuity or intensity systems assures the unreliability of data collected from different institutions. Groups working to advance the Nursing Minimum Data Set are tackling this problem, and perhaps it can be alleviated. If the nursing profession is going to insist on direct calculations of resource consumption from clinical data, however, the reliability issues must be addressed in the development of the systems to accomplish that. Nurse researchers will have to take this on. For sound economic reasons, vendors are conservative about developing radically new products, especially when existing products, such as acuity systems, are selling well. Collaborative research joining nurses with health care economists, informaticians, and others can create prototypes that download clinical data from the record of care given to calculate the costs. As these prototypes are developed and submitted to nursing's scientific community for review, their relative strengths and weaknesses should be scrutinized as a basis for proposing standards for methodology and data. After a second generation of prototypes adhering to the standards has been tested, vendors are more likely to be ready to join with developers for commercial production.

Testing the Utility of the Data Elements for Effectiveness Research

In developing nursing data elements for large databases to be used in effectiveness research, researchers should assess the value of any specific element according to a hierarchy of three criteria. The most basic criterion is validity. Does the item accurately represent a legitimate nursing construct? If not, it has no place in research to

assess nursing effectiveness. The second level is reliability. Even if the item seems to be valid, can nurses collect the data accurately and precisely? If not, the item is useless. The highest level is utility. Even if the item represents a legitimate nursing issue, even if nurses can collect the data accurately, does the item tell us anything about the effectiveness or the costs of nursing? If not, it need not be included in the databases for effectiveness research.

The utility question, like the validity and reliability questions, must be resolved empirically. After selecting data items demonstrated to be valid and reliable, investigators can make pilot tests of their utility by collecting and examining the data from one or a few institutions. Is there variation among patients in the occurrence of the diagnoses, or in the type of intervention used to respond to a given diagnosis, or in the patient outcomes, or in the measures of resource consumption? If the data analysis yields negative answers for any items, those items may be excluded from the database. They may be constants, not variables, or they may represent conditions that nursing does not, after all, influence. At the very least, they require further investigation before they may be incorporated into the database.

Recommendations

The most important point made here is that, however urgently nursing may require inclusion of its data in large databases for effectiveness research, the development and selection of the data items must be as soundly scientific as the subsequent research using the data. Rushing to consensus without adequate development and testing would risk invalidating the very research we hope to support. Instead, we should call on the National Center for Nursing Research to offer research grants to investigators in a variety of sites to carry out the validity, reliability, and utility studies suggested here, including the

development of prototypes. In addition, periodic meetings of those investigators and other invited experts, together with personnel of the National Center for Nursing Research, would enable scholars in this endeavor to discuss progress, share findings, and make recommendations for further research and development. Consensus on items for inclusion in databases for effectiveness research could thus be founded not on opinion or ideology or the persuasive powers of selected leaders, but on research and scientific debate. Initiated immediately, this process could lead us to a valid, reliable, useful nursing data set in five years' time. Although that may seem a long delay to meet an urgent need, the quality of the data set is surely more important than speed. The consequences of the process are too important to permit haste. Research using the databases will inform policy decisions that profoundly affect the practice of nursing and the nation's health.

References

Bulechek, G., & McCloskey, J. (1985). *Nursing interventions: Treatment for nursing diagnoses.* Philadelphia: Saunders.

Bulechek, G. M., & McCloskey, J. C. (1987). Nursing interventions: What they are and how to choose them. *Holistic Nursing Practice, 1*(3), 36-44.

Donabedian, A. (1966). Evaluating the quality of medical care. *Milbank Memorial Fund Quarterly, 44* (3—Supplement), 166-206.

Campbell, C. (1984). Appendix 2: coding diagnoses and interventions for processing. In C. Campbell, *Nursing diagnosis and intervention in nursing practice* (pp. 1823-1850). New York: John Wiley and Sons.

Carroll:Johnson, R. M. (1989). *Classification of nursing diagnosis: Proceedings of the Eighth Conference*. Philadelphia: Lippincott.

Fillebaum, G. G., & Smythe, M. A. (1981). The development, validity, and reliability of the OARS Multidimensional Functional Assessment Questionnaire. *Journal of Gerontology, 36*, 428-434.

Grobe, S. (1989). Nursing intervention lexicon and taxonomy study: Description of a conceptual framework. In B. Barker, D. Cao, D. Quin, & G. Wagner (Eds.), *Medinfo 89* (pp. 925-927). Amsterdam: North Holland.

Hastings, C., & Muir-Nash, J. (1988). Validation of a taxonomy of ambulatory nursing practice. *Nursing Economics, 7*(3), 142-149.

Horn, B. J., & Swain, M. A. (1977). *Development of criterion measures of nursing care (vols. 1 and 2).* (Publication nos. PB267004/LP and PB267005/LP). Springfield, VA: National Technical Information Service.

Joel, L. A. (1984). DRGs and RIMs: Implications for nursing. *Nursing Outlook, 32*(1), 42-49.

Kim, M. J., McFarland, G. K., & McLane, A. M. (1984). Approved nursing diagnoses and small group work on diagnostic labels. In M. J. Kim, G. K. McFarland, & A. M. McLane (Eds.), *Classification of nursing diagnoses: Proceedings of the Fifth National Conference* (pp. 470-477). St Louis: Mosby.

Martin, K., Scheet, N., Crews, C., & Simmons, D. (1986). *Client Management Information System for Community Health Nursing Agencies: An Implementation Manual.* Rockville, MD: Division of Nursing, USDHHS.

Simmons, D. (1980). *A classification scheme for client problems in community health nursing. Nurse planning information series, Vol. 14.* (Publication no. [HRP]-0501501). Springfield, VA: National Technical Information Service.

Ver Steeg, D. F. (1988). Computer use and nursing research: A taxonomy of nursing diagnoses for use in ambulatory care nursing education, practice, and research. Part I: Development and major headings. *Western Journal of Nursing Research, 10*(4), 506-509.

Waltz, C. F., & Strickland, O. L. (Eds.). (1988). *Measurement of nursing outcomes. Volume One: Measuring client outcomes.* New York: Springer.

Visiting Nurse Association of Omaha. (1986). Client management information system for community health nursing agencies. (Publication no. 0907-023). Springfield, VA: National Technical Information Service.

Werley, H. H., & Lang, N. M. (Eds.) (1988). *Identification of the nursing minimum data set.* New York: Springer.

Werley, H. H., & Lang, N. M. (1988). The consensually derived nursing minimum data set: Elements and definitions. In H. H. Werley & N. M. Lang (Eds.), *Identification of the nursing minimum data set.* New York: Springer.

Zimmer, M. J. (Ed.). (1974). Symposium on quality assurance. *Nursing Clinics of North America, 9* (2), 303-379.