Outcomes Research

Chapter Overview

The Theoretical Basis of Outcomes Research 273
Donabedian’s Theory of Quality Health Care 273
The Agency for Health Services Research 281
The Agency for Health Care Policy and Research 281
Outcomes Research and Nursing Practice 286
The American Nurses Association’s “Nursing’s Safety & Quality Initiative” 286
Nursing-Sensitive Patient Outcomes 290

Learning Outcomes

After completing this chapter, you should be able to:

1. Discriminate between traditional quantitative research and outcomes research.
2. Explain the theoretical basis of outcomes research.
3. Explain the importance of outcomes research.
4. Identify nursing-sensitive patient outcomes.
5. Describe some of the unique methodologies of outcomes research.

Key Terms

Clinical decision analysis, p. 302
Consensus knowledge building, p. 298
Cost-benefit analysis, p. 306
Cost-effectiveness analysis, p. 305
Costs, p. 306
Design, p. 298
Efficiency, p. 305
Geographical analyses, p. 305
Health, p. 273
Interdisciplinary teams, p. 304
Intermediate end points, p. 307
Measurement error, p. 297
Medical, p. 283
Multilevel analysis, p. 315
Opportunity costs, p. 306
Outcomes research, p. 273
Out-of-pocket costs, p. 306
Population-based studies, p. 302
Practice pattern profiling, p. 299
Prospective cohort study, p. 301
Quality, p. 273
Research tradition, p. 294
Retrospective cohort study, p. 301
Sampling error, p. 297
Small area analyses, p. 305
Standard of care, p. 276
Standardized mortality ratio (SMR), p. 301
Structures of care, p. 277
Variance analysis, p. 313
Outcomes research focuses on the end results of patient care. In order to explain the end results, nurse researchers also must understand the processes used to provide patient care. The strategies used in outcomes research are, to some extent, a departure from the accepted scientific methodology for health care research, and they incorporate evaluation methods, epidemiology, and economic theory. The findings of outcome studies continue to have a powerful impact on the provision of health care and the development of health policy.

The momentum propelling outcomes research comes not from scholars but from policy makers, insurers, and the public. These groups increasingly demand that providers justify interventions and systems of care in terms of improved patient lives and that costs of care be considered in the evaluation of treatment outcomes (Hinshaw, 1992). A major shift has occurred in published nursing studies, with the number of studies using traditional quantitative or qualitative methods being dwarfed by the number of outcomes studies. A large number of nursing and multidisciplinary journals focused on outcomes research have been initiated. A listing of these journals is presented in Table 9-1.

This chapter provides an explanation of the theoretical basis of outcomes research, a brief history of the emerging endeavors to examine outcomes, the importance of outcomes research designed to examine nursing practice, and methodologies used in outcomes research.

The Theoretical Basis of Outcomes Research


Donabedian’s Theory of Quality Health Care

Quality is the overriding construct of Donabedian’s Theory of Quality Health Care, although he never defines this concept (Mark, 1995). The cube shown in Figure 9-1 helps explain the elements of quality health care. The three dimensions of the cube are health, subjects of care, and providers of care. The concept health has many aspects; three are shown on the cube: physical-physiological function, psychological function, and social function. Donabedian (1987, p. 4) proposes that “the manner in which we conceive of health and of our responsibility for it, makes a fundamental difference to the concept of quality and, as a result, to the methods that we use to assess and assure the quality of care.”

Loegering, Reiter, and Gambone (1994) modified Donabedian’s levels to include the patient, family, and community as providers of care as well as recipients of care. They suggest that access to care is one dimension of the provision of care by the community. Figure 9-2 illustrates their modifications.
Donabedian (1987) identifies three objects of evaluation in appraising quality: structure, process, and outcome. A complete quality assessment program requires the simultaneous use of all three concepts and an examination of the relationships among the three. However, researchers have had little success in accomplishing this theoretical goal. Studies designed to examine all three concepts would require sufficiently large samples of various structures, each with the various processes being compared and large samples of subjects who have experienced the outcomes of those processes. The funding and the cooperation necessary to accomplish this goal are not yet available.

Evaluating Outcomes

The goal of outcomes research, the evaluation of outcomes as defined by Donabedian, is not as simplistic as it might immediately appear. Donabedian’s theory requires that identified outcomes be clearly linked with the process that caused the outcome. To accomplish this linking, the researcher must define the process and justify the causal links with the selected outcomes. The identification of desirable outcomes requires dialogue between the subjects of care and the providers of care. Although the providers of care may delineate what is

<table>
<thead>
<tr>
<th>Table 9–1</th>
<th>Nursing and Multidisciplinary Journals Focused on Outcomes Research</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best Practice</td>
<td>Clinical Effectiveness in Nursing</td>
</tr>
<tr>
<td>Clinical Effectiveness in Nursing</td>
<td>Effective Health Care</td>
</tr>
<tr>
<td>Effective Health Care</td>
<td>Evidence-Based Mental Health</td>
</tr>
<tr>
<td>Evidence-Based Practice</td>
<td>Health Education Research</td>
</tr>
<tr>
<td>Health Education Research</td>
<td>Health Promotion Practice</td>
</tr>
<tr>
<td>Health Promotion Practice</td>
<td>Health Technology Assessment</td>
</tr>
<tr>
<td>Health Technology Assessment</td>
<td>International Journal of Health Care Quality Assurance</td>
</tr>
<tr>
<td>International Journal of Health Care Quality Assurance</td>
<td>Incorporating Leadership in Health Services</td>
</tr>
<tr>
<td>Journal of Care Management</td>
<td>Journal of Clinical Ethics</td>
</tr>
<tr>
<td>Journal of Clinical Ethics</td>
<td>Journal of Clinical Outcomes Management</td>
</tr>
<tr>
<td>Journal of Clinical Outcomes Management</td>
<td>Journal for Health Care Quality</td>
</tr>
<tr>
<td>Journal for Health Care Quality</td>
<td>Journal of Health Services and Research Policy</td>
</tr>
<tr>
<td>Journal of Health Services and Research Policy</td>
<td>Journal of Integrated Care Pathways</td>
</tr>
<tr>
<td>Journal of Integrated Care Pathways</td>
<td>Journal of Nursing Care Quality</td>
</tr>
<tr>
<td>Journal of Nursing Care Quality</td>
<td>Journal of Nursing Quality Assurance</td>
</tr>
<tr>
<td>Journal of Nursing Quality Assurance</td>
<td>Journal of Quality in Clinical Practice</td>
</tr>
<tr>
<td>Journal of Quality in Clinical Practice</td>
<td>Journal of Rehabilitation Outcomes Measurement</td>
</tr>
<tr>
<td>Journal of Rehabilitation Outcomes Measurement</td>
<td>Nursing and Health Policy Review</td>
</tr>
<tr>
<td>Nursing and Health Policy Review</td>
<td>Outcomes Management</td>
</tr>
<tr>
<td>Outcomes Management</td>
<td>Outcomes Management for Nursing Practice</td>
</tr>
<tr>
<td>Outcomes Management for Nursing Practice</td>
<td>Patient Satisfaction Management</td>
</tr>
<tr>
<td>Patient Satisfaction Management</td>
<td>Quality in Health Care</td>
</tr>
<tr>
<td>Quality in Health Care</td>
<td>Quality in Primary Care</td>
</tr>
<tr>
<td>Quality in Primary Care</td>
<td>Quality and Safety in Health Care</td>
</tr>
<tr>
<td>Quality and Safety in Health Care</td>
<td>Quality Management in Health Care</td>
</tr>
<tr>
<td>Quality Management in Health Care</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER 9
Outcomes Research

Figure 9-1. Level and scope of concern as factors in the definition of quality. (From Donabedian, A. [1987]. Some basic issues in evaluating the quality of health care. In L. I. Rinke [Ed.], Outcome measures in home care [Vol. 1, pp. 3–28]. New York: National League for Nursing.)

Figure 9-2. Various levels at which the quality of health care can be assessed. (From Donabedian, A. [1988]. The quality of care: How can it be assessed? Journal of the American Medical Association, 260[12], 1744. Copyright 1988, American Medical Association.)
achievable, the subjects of care must clarify what is desirable. The outcomes also must be relevant to the goals of the health care professionals, the health care system of which the professionals are a part, and society.

Outcomes are time dependent. Some outcomes may not be apparent for a long period after the process that is purported to cause them, whereas others may be apparent immediately. Some outcomes are temporary and others are permanent. Thus, the selection of an appropriate time frame for determining the selected outcomes must be established.

A final obstacle to outcomes evaluation is one of attribution. This requires assigning the place and degree of responsibility for the outcomes observed. A specific outcome often is influenced by a multiplicity of factors. Lewis (1995) points out that health care represents only one dimension of a complex situation. Patient factors, such as compliance, predisposition to disease, age, propensity to use resources, high-risk behaviors (e.g., smoking), and lifestyle, also must be taken into account. Environmental factors such as air quality, public policies related to smoking, and occupational hazards also must be included. Responsibility for outcomes may be distributed among providers, patients, employers, insurers, and the community.

As yet, little scientific basis has been established for judging the precise relationship between each of these factors and the selected outcome. Many of the influencing factors may be outside the jurisdiction or influence of the health care system or of the providers within it. One solution to this problem of identifying relevant outcomes is to define a set of closely related outcomes specific to the condition for which care is being provided. Critical pathways and care maps may be useful in defining at least related outcomes. However, related outcomes do not provide the degree of evidence of examining the desired outcomes.

Evaluating Process

The process of clinical management has been, for most health care professionals, an art rather than a science. Understanding the process sufficiently to study it must begin with much careful reflection, dialogue, and observation. Clinical management has multiple components, many of which have not yet been clearly defined or tested. Bergmark and Oscarsson (1991, pp. 139–140) suggest the following questions as important to consider in evaluating process: (1) “What constitutes the ‘therapeutic agent’?” (2) “Do practitioners actually do what they say they do?” (3) “Do practitioners always know what they do?” Current outcomes studies are using process variables that are easy to identify. Answers to questions such as those posed by Bergmark and Oscarsson are more difficult to define and initially will require observation, interviews, and the use of qualitative research methodologies.

Three components of process of particular interest to Donabedian are standards of care, practice styles, and costs of care.

Standards of Care

A standard of care is a norm by which quality of care is judged. Clinical guidelines, critical paths, and care maps define standards of care. According to Donabedian (1987), a practitioner has legitimate responsibility to apply available knowledge in the management of a dysfunctional state. This management consists of (1) the identification or diagnosis of the
dysfunction, (2) the decision whether to intervene, (3) the choice of intervention objectives, (4) the choice of methods and techniques to achieve the objectives, and (5) the skillful execution of the selected techniques.

Donabedian (1987) recommends the development of criteria to be used as a basis for judging the quality of care. These criteria may take the form of clinical guidelines or care maps based on previous validation of the contribution of the care to outcomes. The clinical guidelines published by the Agency for Healthcare Research and Quality (AHRQ) establish norms on which the validity of clinical management can be judged. However, the core of the problem, from Donabedian’s perspective, is clinical judgment. Analysis of the process of making diagnoses and therapeutic decisions is critical to the evaluation of the quality of care. The emergence of decision trees and algorithms is a response to Donabedian’s concerns and provides a means of evaluating the adequacy of clinical judgments.

Practice Styles

The style of practice is another dimension of the process of care that influences quality; however, it is problematic to judge what constitutes “goodness” in style and to provide justification for the decisions. Moreover, diverse styles of interpersonal relationships are possible. Most studies examining practice styles are conducted with physicians as subjects. A few studies of practice styles of nurses, however, are beginning to appear in the literature (Bircumshaw & Chapman, 1988; Fullerton, Hollenbach, & Wingard, 1996).

Costs of Care

A third dimension of the examination of quality of care is cost. Maintaining a specified level of quality of care necessarily has cost consequences. Providing more and better care is likely to increase costs but also is likely to produce savings. Economic benefits can be obtained by preventing illness, preventing complications, maintaining a higher quality of life, or prolonging productive life.

A related issue is who bears the costs of care. Some measures purported to reduce costs have instead simply shifted costs to another party. For example, in certain instances a hospital can reduce its costs by discharging a particular type of patient early, but total costs will increase if the necessary community-based health care raises costs above those incurred by keeping the patient hospitalized longer. In this case, the third-party provider may experience higher costs. In many cases, the costs are shifted from the health care system to the family as out-of-pocket costs. Studies examining changes in costs of care must consider total costs.

Evaluating Structure

Structures of care are the elements of organization and administration that guide the processes of care. The first step in evaluating structure is to identify and describe the elements of the structure. Various administration and management theories may be used to identify the elements of structure to be studied. Examples of such elements are leadership, tolerance of innovativeness, organizational hierarchy, decision-making processes, distribution of power, financial management, and administrative decision-making processes.
The second step is to evaluate the impact of various structure elements on the process of care and on outcomes. This evaluation requires comparing different structures that provide the same processes of care. In the evaluation of structures, the unit of measure is the structure. The evaluation requires access to a sufficiently large sample of like structures with similar processes and outcomes, which can then be compared with a sample of another structure providing the same processes and outcomes. For example, a researcher may decide to compare various structures providing primary health care, such as the private physician office, the health maintenance organization (HMO), the rural health clinic, the community-oriented primary care clinic, and the nurse-managed center. Another researcher may wish to examine surgical care provided within the structures of a private outpatient surgical clinic, a private hospital, a county hospital, and a teaching hospital associated with a health science center. In each of these studies, the focus will be the impact of structure on processes of care and outcomes of care. See Figures 9-3 through 9-6 for examples of frameworks for outcomes studies with this focus.

**Figure 9-3.** The conceptual model: Economic, clinical, and humanistic outcome (ECHO) model. (From Kozma, C. M., Reeder, C. E., & Schulz, R. M. [1993]. Economic, clinical, and humanistic outcomes: A planning model for pharmacoeconomic research. *Clinical Therapeutics, 15*(6), 1125.)
CHAPTER 9 Outcomes Research

Figure 9-4. A systems perspective of health services research. (From Anderson, R. M., Davidson, P. L., & Ganz, P. A. [1994]. Symbiotic relationships of quality of life, health services research and other health research. Quality of Life Research, 3[5], 367.)

Figure 9-5. The health services system. (From Vivier, P. M., Bernier, J. A., & Starfield, B. [1994]. Current approaches to measuring health outcomes in pediatric research. Current Opinions in Pediatrics, 6[5], 531.)
When critiquing an outcomes research study, consider structure of care. Then ask yourself the following questions:

1. Do the outcome variables show a clear link between the process of care and the identified outcome?
2. Does the desirability of the selected outcome reflect the preference of the patient, rather than the provider?
3. Is the process of care leading to the outcome clearly defined?
4. The process of care is driven by the structure of care. Is the structure of care defined in the study?
5. The practice style of providers is a major factor in the process of care. Is the practice style defined?

**Figure 9-6.** Relationships among measures of patient outcome in a health-related quality-of-life conceptual model. (From Wilson, I. B., & Cleary, P. D. [1995]. Linking clinical variables with health-related quality of life: A conceptual model of patient outcomes. *Journal of the American Medical Association, 273*(1), 60. Copyright 1995, the American Medical Association.)
The Agency for Health Services Research

Nurses participated in the initial federal involvement in the quality of health care. In 1959, two National Institutes of Health Study Sections, the Hospital and Medical Facilities Study Section and the Nursing Study Section, met to discuss concerns about the adequacy and appropriateness of medical care, patient care, and hospital and medical facilities. As a result of their dialogue, a Health Services Research Study Section was initiated. This study section eventually became the Agency for Health Services Research (AHSR). With small amounts of funding from Congress, the AHSR continued to study the effectiveness of health services, primarily supporting the research of economists, epidemiologists, and health policy analysts (White, 1993). Two projects that were to have the greatest impact were small area analyses and the Medical Outcomes Study (MOS).

Small Area Analyses

In the 1970s, an epidemiologist named Wennberg began a series of studies examining small area variations in medical practice across towns and counties. He found a wide variation in the tonsillectomy rate from one town to another in the New England area that could not be explained by differences such as health status, insurance, and demographics. These findings were replicated for a variety of medical procedures. Investigators began a search for the underlying causes of such variation and their implications for health status (O’Connor, Plume, & Wennberg, 1993; Wennberg, Barry, Fowler, & Mulley, 1993). Studies also revealed that many procedures, such as coronary artery bypass, were being performed on patients who did not have appropriate clinical indications for such surgery (Power, Tunis, & Wagner, 1994).

The Medical Outcomes Study

The Medical Outcomes Study (MOS) was the first large-scale study to examine factors influencing patient outcomes. The study was designed to identify elements of physician care associated with favorable patient outcomes. The conceptual framework for the MOS is shown in Figure 9-7. Variations in use of resources and in physician technical and interpersonal styles were examined (Greenfield et al., 1992; Kelly, Huber, Johnson, McCloskey, & Maas, 1994; Riesenberg & Glass, 1989; Stewart et al., 1989). Kelly and colleagues noted, however, that the MOS failed to control for the effects of nursing interventions, staffing patterns, and nursing practice delivery models on medical outcomes. Coordination of care, counseling, and referrals—activities more commonly performed by nurses than by physicians—were considered in the MOS to be components of medical practice.

The Agency for Health Care Policy and Research

The Agency for Health Care Policy and Research (AHCPR), created in 1989 by Congress, replaced the AHSR. The congressional mandate for the AHCPR was

…to carry out research, demonstrations, guideline development, training, and dissemination activities with respect to health care services and systems of information regarding the following
areas: the effectiveness, efficiency, quality, and outcomes of health services; clinical practice, including primary care; health care technologies, facilities, and equipment; health care costs, productivity, and market forces; health promotion and disease prevention; health statistics and epidemiology; and medical liability. (Gray, 1992, p. 40)

A National Advisory Council for Health Care Policy, Research, and Evaluation also was established by Congress. The Council was required to include (1) health care researchers; (2) health care professionals (specifically including nurses); (3) professionals from the fields of business, law, ethics, economics, and public policy; and (4) persons representing the interests of consumers. The budget for the AHCPR increased to $1.9 million in 1988, $5.9 million in 1989, and $37.5 million in 1990.

The AHCPR initiated several major research efforts to examine medical outcomes. Two of the most significant, described next, are the Medical Treatment Effectiveness Program (MEDTEP) and a component of MEDTEP referred to as Patient Outcomes Research Teams (PORTs) (Greene, Bondy, & Maklan, 1994).

The Medical Treatment Effectiveness Program

MEDTEP was established by Congress in 1989 to be implemented by the AHCPR. The purpose of the program was to improve the effectiveness and appropriateness of medical practice. The term medical was used by Congress when the program was mandated. However, it was broadly interpreted to include health care in general and nursing care in particular—an important consideration from the perspective of the nursing profession. The program was charged to develop and disseminate scientific information about the effects of health care services and procedures on patients’ survival, health status, functional capacity, and quality of life, a remarkable shift from the narrow focus of traditional medical research. MEDTEP funded three research areas: (1) patient outcomes research, (2) database development, and (3) research on effective methods of disseminating the information gathered. In 1993, the program implemented studies to examine the effects of pharmaceuticals on patient outcomes and provided $19 million to establish Research Centers on Minority Populations (Clinton, 1993).

Patient Outcomes Research Team Projects

In 1994, Congress mandated patient outcomes research team projects (PORTs), that is, large-scale, multifaceted, and multidisciplinary projects to “identify and analyze the outcomes and costs of current alternative practice patterns in order to determine the best treatment strategy and to develop and test methods for reducing inappropriate variations” (U.S. Congress, 1994, p. 67). The PORTs were required to “conduct literature reviews and syntheses; analyze practice variations and associated patient outcomes, using available data augmented by primary data collection where desired; disseminate research findings; and evaluate the effects of dissemination” (U.S. Congress, 1994, p. 67). Questions typically addressed by PORTs include the following:

- Do patients benefit from the care provided?
- What treatments work best?
- Has the patient’s functional status improved?
- From whose viewpoint has the patient’s outcome been evaluated?
- Are health care resources well spent? (Tanenbaum, 1994; Wood, 1990)

A major task of PORTs was to disseminate their findings and change the practice of health care providers to improve patient outcomes. A framework for dissemination was developed that identified the audiences for disseminated products, the media involved, and the strategies that foster assimilation and adoption of information (Goldberg, Cummings, Steinberg, Ricci, Shannon, Soumerai, et al., 1994).

A Cost of Care Workgroup, consisting of a representative from each PORT, was convened in 1994 with the following four goals: (1) to determine the best methods for estimating the cost of certain conditions using claims data, (2) to evaluate methods for estimating the cost of care using billing information and patient interview data, (3) to examine methods for determining the indirect cost of care, and (4) to evaluate methods for comparing the cost of care internationally (Lave et al., 1994).

In 1992, the National Center for Nursing Research (NCNR) sponsored a Conference on Patient Outcomes Research: Examining the Effectiveness of Nursing Practice. In the keynote speech, Hinshaw, then director of the NCNR, made the following suggestions:
From a nursing perspective, particular clinical conditions need to be identified that are more specific to nursing’s focus on prevention, health promotion, symptom management, and the amelioration of the effects of acute and chronic illnesses. We are all familiar with clinical conditions that are central to our practice, such as skin integrity, pain, urinary incontinence, nausea and vomiting, nutritional deficits, confusion, restricted mobility, depression, fatigue, and illness-related stress. It will be particularly important in our research programs that we begin to both define and refine the patient outcomes specific to interventions focused on such clinical conditions. (Hinshaw, 1992, p. 9)

Examining the impact of nursing on overall hospital outcomes requires inclusion of nursing data in the large databases used to analyze outcomes. The cost of adding new variables to these databases is high. Nursing professionals are competing with the voices of others who wish to add their own relevant variables. However, with the force of the American Nurses Association (ANA), the voice of the profession is being heard.

The NCNR, now the National Institute for Nursing Research (NINR), developed a partnership with the AHCPR to fund outcomes studies of importance to nursing. Calls for proposals jointly supported by the AHCPR and the NINR are announced each year. (These calls for proposals can be found on the NINR home page on the World Wide Web: http://www.nih.gov/ninr/.)

With a growing budget and strong political support, proponents of the AHCPR were becoming a powerful force demanding change in health care, because of the demand for health care reform that existed throughout the government and among the public. The role of the AHCPR, however, was the focus of considerable controversy. For example, a subset of spinal orthopedic surgeons strongly opposed the guidelines for the treatment of back pain. From 1994-1998, this group of physicians assailed members of Congress with visits, letters, and telephone calls. The AHCPR was under attack by powerful forces and was in danger of being eliminated or of having its budget greatly reduced. In fact, for the 1997 fiscal year budget, Congress cut AHCPR funding by $35 million, thereby ending the funding of MEDTEP and new PORTs. The following year, however, Congressional members recognized that the opposition to the agency had been based on special interests and that the contested PORT findings were valid (Deyo, Psaty, Simon, Wagner, & Omenn, 1997; Fardon, Garfin, & Saal, 1997).

The intimidation of agencies and researchers by special-interest groups has serious implications for both scientists and society. Fardon, Garfin, and Saal, in a letter to The New England Journal of Medicine (1997), express concern about this problem.

Harassment of researchers and funding agencies is a substantial disincentive to pursuing certain research on medical care or health risks. In effect, special-interest groups with money and power want to define acceptable questions and shape the range of acceptable answers. Eliminating public, peer-reviewed funding would slow the production of objective knowledge, force investigators to seek funding that may not be free of conflict of interest, and leave patients, physicians, and insurers without essential scientific evidence. University faculty members are governed by financial conflict-of-interest rules intended to prevent them from conducting research in which they or their relatives might have a financial stake. Thus, the elimination of public research support and the intimidation of independent investigators are inimical to larger social interests. Professional societies, universities, and the government need to weigh in quickly and heavily against strategies and specific cases of intimidation and vengeful budget cuts.... Inquiry may be warranted concerning the extent to which special-interest groups block or delay the publication...
of unwanted findings. Journals may need to make a special effort to avoid relying on otherwise highly qualified reviewers and editorialists who have financial conflicts of interests, especially consultants to firms whose products receive negative evaluations. Journals may also need to set up defenses against potential threats of withholding advertising. When funding agencies come under attack from groups with narrow interests, prompt and unambiguous responses from universities and professional organizations are needed. Self-interested attacks must be pointed out to politicians, who may otherwise be unable to distinguish self-interested parties from disinterested ones (Deyo et al., 1997, pp. 1315-1316).

The AHCPR operated without authorization from 1995 until December 6, 1999, receiving operating funds through congressional appropriations. The reauthorization act changed the name of the AHCPR to the Agency for Healthcare Research and Quality (AHRQ). The AHRQ is designated as a scientific research agency. The term policy was removed from the agency name, to avoid the perception that this body determined federal health care policies and regulations. The word quality was added to the agency's name, establishing the AHRQ as the lead federal agency on quality of care research, with a new responsibility to coordinate all federal quality improvement efforts and health services research. The new legislation eliminated the requirement that the AHRQ develop clinical practice guidelines. However, the AHRQ still supports these efforts through Evidence-Based Practice Centers and the dissemination of evidence-based guidelines through its National Guideline Clearinghouse. The new legislation defines the AHRQ's mission as follows:

- Meet the information needs of its customers—patients and clinicians, health system leaders, and policymakers—so that they can make more informed healthcare decisions.
- Build the evidence base for what works and doesn't work in healthcare and develop the information, tools, and strategies that decisionmakers can use to make good decisions and provide high-quality healthcare based on evidence.
- Develop scientific knowledge in these areas but will not mandate guidelines or standards for measuring quality. (One Hundred Sixth Congress of the United States, 1999, pp. 2–3)

The proposed budget for the AHRQ in fiscal year 2004 was $279 million.

The United States is not the only country making demands for improvements in quality of care and reductions in costs. Many countries around the world are experiencing similar concerns and addressing them in relation to their particular government structure. Thus, the movement into outcomes research and the application of the approaches described in this chapter constitute a worldwide phenomenon.

Reviews of Evidence from Existing Knowledge

The extensive review of published and unpublished work related to a particular health care problem was taken over by a number of organizations located in various countries. The most well known of these is the Cochrane Collaboration located in England. The Cochrane Collection is available on the Internet and through some search engines such as CINAHL. Such reviews may conclude that available knowledge is insufficient for the development of clinical guidelines, or they may propose clinical guidelines and/or a protocol developed from the material reviewed. Reviews are evaluated regularly, and a judgment is made about the need to update the review. The date of the initial review and all updates are provided.
Clinical Guideline Panels

Clinical guideline panels are developed to incorporate available evidence on health outcomes into sets of recommendations concerning appropriate management strategies for patients with the studied conditions. Any professional group may gather a group to develop guidelines on a particular topic. Some groups seek funding for the project, whereas others, such as professional organizations, conduct the work as an aspect of the organizational work. Medical schools and nursing schools have submitted guidelines as have medical and nursing organizations and volunteer agencies such as the American Cancer Society. Guidelines developed across the world are included. Some guidelines are evidence-based, whereas others are not. The evidence-based guidelines have considerably more validity. (Current guidelines can be obtained from the National Guideline Clearinghouse of AHRQ at the following Internet address: http://www.guideline.gov/.)

Outcomes Research and Nursing Practice

Outcome studies provide rich opportunities to build a stronger scientific underpinning for nursing practice (Rettig, 1990): “Nursing needs to be able to explain the impact of care provided by its practitioners through measures of outcomes of patient care that reflect nursing practice” (Moritz, 1991, p. 113).

The American Nurses Association’s “Nursing’s Safety & Quality Initiative”

In the late 1980s and early 1990s, hospitals were confronted with managed care requirements to reduce costs. To accomplish this goal, hospitals across the country reduced their nursing staff and replaced them with unlicensed personnel having very little training for their assignments. Managed care dictated earlier patient discharges, resulting in patient loads that were sicker and required complex care. Nurses repeatedly complained that patient care was inadequate and that patients were experiencing complications and dying needlessly because of the inadequate staffing of RNs. However, nursing had little concrete evidence of these statements. Many nurses left nursing practice or changed to community areas of nursing practice. The RNs remaining in hospital practice tended to be new graduates who were placed in positions of responsibility without adequate experience. Assignment loads became increasingly difficult. Nurses tended to leave the hospital after 2 years, and more new graduates replaced them. A shortage of nurses, already in place, was exacerbated by this situation. Recruitment of new students became increasingly difficult as greater numbers of news items discussed the problems.

In 1994 the ANA, in collaboration with the American Academy of Nursing Expert Panel on Quality Health Care (Mitchell, Ferketich, & Jennings, 1998), launched an initiative to identify indicators of quality nursing practice, and to collect and analyze data using these indicators across the United States. The goal was to identify or develop nursing-sensitive quality measures. Donabedian’s theory was used as the framework for the project. The committee conducted an extensive literature review, expert panel discussions, and focus group interviews to identify 21 nursing care indicators relevant to nursing care quality in acute care. On the basis of Donabedian’s theory or established evidence of a strong link to nursing care quality, the committee then reduced the list of indicators to 10. The 10 indicators currently being collected and 4 others undergoing pilot testing are listed in Table 9-2.
If nursing care could be compared among hospitals, nursing would have the evidence to justify claims of patient harms from the changes in RN staffing. ANA entered a new area of research and asked questions that had not previously been studied. No one knew what indicators were sensitive to the nursing care provided to patients or what relationships existed between nursing inputs and patient outcomes. They had to persuade hospitals to participate in the study at a time when hospitals had a severe case of “data paranoia” (fear of providing data to anyone outside the hospital because of how third parties might interpret the data). Every hospital had a different way of measuring the indicators selected by ANA. Persuading them to change to a standardized measure of the indicators for consistency across hospitals was a major endeavor (Jennings, Loan, DePaul, Brosch, & Hildreth, 2001; Rowell, 2001).

Nurse researchers and cooperating hospitals instituted the mechanisms required for data collection and began multiple pilot studies. These pilot studies identified multiple obstacles to the project. They learned that not only must the indicators be measured consistently, but that data collection must be standardized. As studies continued, indicators were amplified and continue to be tested. As this testing continues, further alterations in the indicators occur (Anonymous, 1997; Campbell-Heider, Krainovich-Miller, King, Sedhom, & Malinski, 1998; Jennings et al., 2001).

The ANA proposes that all hospitals collect and report on the 10 nursing-sensitive quality indicators. The ANA is working to ensure that these indicators are included in data collected by accrediting organizations and by the federal government, and that the data be shared with key groups. The ANA also is encouraging state nurses’ associations to lobby state legislatures to include the nursing-sensitive quality indicators into regulations or state law.

In 1998, the ANA provided funding to develop a national database to house data collected using nursing-sensitive quality indicators. This database, named The National Database of Nursing Quality Indicators (NDNQI), is a program of the National Center for

---

**Table 9-2 American Nurses Association Acute Care Nursing-Sensitive Quality Indicators**

<table>
<thead>
<tr>
<th>Indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient falls</td>
</tr>
<tr>
<td>Patient falls with injury</td>
</tr>
<tr>
<td>Pressure ulcers—% of patients with documented ulcer (stage I–IV) on day of prevalence study. Also have Hospital-acquired ulcer—% of patients with documented ulcer (stage I–IV) on day of prevalence study</td>
</tr>
<tr>
<td>Nurse satisfaction</td>
</tr>
<tr>
<td>Nursing Hours Per Patient Day (HPPD)—RN, LPN/LVN, UAP—number of productive hours worked by nursing staff with direct patient care responsibilities</td>
</tr>
<tr>
<td>Staff mix—the total number of productive hours worked by each skill mix category (RN, LPN, UAP)/total staff hours</td>
</tr>
<tr>
<td>Type of unit (critical care, step down, medical, surgical and combined)</td>
</tr>
<tr>
<td>Number of staffed beds designated by the hospital</td>
</tr>
<tr>
<td>Agency staff—total number of productive hours worked by contract staff</td>
</tr>
<tr>
<td>Urban vs. rural category</td>
</tr>
</tbody>
</table>

**New Indicators Undergoing Pilot Testing**

- Pediatric pain
- Peripheral intravenous infiltration
- Restraint use
- Patient aggression

LPN, licensed practical nurse; LVN, licensed vocational nurse; RN, registered nurse; UAP, unlicensed assistive personnel.

Nursing Quality, funded by the ANA. The database is housed at the University of Kansas Medical Center Research Institute (KUMCRI) and of the University of Kansas School of Nursing. In 2001, data from nursing-sensitive quality indicators were being collected from more than 120 hospitals in 24 states across the United States. By 2005, that number had increased to 767 hospitals in the 50 states and the District of Columbia. The National Center for Nursing Quality analyzes the data quarterly and provides feedback reports to all participating hospitals. Confidential benchmarking reports are provided to allow hospitals to compare their results with those of other hospitals (Rowell, 2001).

In 1997, the ANA appointed members of an Advisory Committee on Community-Based Non-Acute Care Indicators to identify the first core set of indicators for non-acute care settings. Some of the members had helped develop the Acute Care Indicators, giving the committee some continuity of the work. The committee began by selecting a theoretical base for its work: Evans and Stoddart’s (1990) determinants of health model and also Donabedian’s model of quality. As its work progressed, the committee chose to synthesize a model to guide the identification and testing of indicators (Figure 9-8). The committee followed the acute care group’s process in selecting the indicators. A hired contractor conducted the literature review, while the committee conducted focus groups and interviews with key stakeholders such as consumers of care, registered nurses, policy makers, regulators, payers, facility administrators, and purchasers. Owing to budget constraints, the ANA limited development to 10 indicators (see Table 9-3). The committee requests that all nurses and nursing organizations join with the ANA to continue to expand this work (Head, Mass, & Johnson, 2003; Sawyer et al., 2002). (For current information on the ANA’s Safety & Quality Initiative, visit the following website: http://nursingworld.org/quality.)

A number of studies on the effects of nursing staff mix on patient outcomes have been published recently (Blegen, Goode, & Reed, 1998; Buerhaus & Needleman, 2000; Cho, Ketefian, Barkauskas, & Smith, 2003; Hall, Doran, Baker, Pink, Sidani, O’Brien, et al., 2001; Houser, 2000; Needleman, Buerhaus, Mattke, Steward, & Zelevinsky, 2002a, 2002b). These studies are finding a significant effect of staffing mix on patient outcomes.

Sochalski (2001), after a review of studies on staff mix, cautions that

‘missing from these studies is a more thorough explanation of how nurse staffing affects patient outcomes. That is, does increasing staffing levels and/or skill mix, under any circumstances, yield better outcomes, or are the effects of staffing titrated by other features in the practice environment that influence nursing’s ability to deliver the quality of care that results in better patient outcomes? Trying to establish minimum staffing ratios in the absence of clear information on just how staffing levels affect outcomes may result in ratios that overestimate or underestimate what is really needed to improve patient care. Furthermore, if the effect of staffing on patient out-

Table 9–3  Nursing-Sensitive Quality Indicators for Community-Based Non-Acute Care Indicators

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain management (symptom severity)</td>
<td>The treatment and prevention of pain and discomfort. Effectiveness is related to level of functioning and activities of daily living and includes measures of frequency, intensity and duration of pain symptoms.</td>
</tr>
<tr>
<td>Consistency of communication (strength of therapeutic alliance)</td>
<td>Consistent RN/advanced practice registered nurses (APRN) provider identified in the data/record.</td>
</tr>
<tr>
<td>Staff mix (utilization of services)</td>
<td>Total number of direct care hours or total number of encounters provided by RN or APRN staff who have client care responsibilities (per episode/encounter/case as appropriate to the setting) and (RNs, LPNs, UAPs caring for clients) – The percent of registered nursing care hours as a total of all nursing care hours; secondary measure—percent of APRNs.</td>
</tr>
<tr>
<td>Client satisfaction</td>
<td>The degree to which the care received met client expectations regarding nursing care, pain management, patient education and overall care.</td>
</tr>
<tr>
<td>Prevention of tobacco use (risk reduction)</td>
<td>The number of clients attending educational sessions per year provided and/or coordinated by RNs about the risks of tobacco use (includes: coordination of educational sessions/programs either with individuals or groups).</td>
</tr>
<tr>
<td>Cardiovascular prevention (risk reduction)</td>
<td>The number of clients attending educational sessions per year provided and/or coordinated by RNs about risks of cardiovascular disease.</td>
</tr>
<tr>
<td>Care giver activity (protective factors)</td>
<td>The existence or frequency of primary care giver involvement.</td>
</tr>
<tr>
<td>Identification of primary care giver (protective factors)</td>
<td></td>
</tr>
<tr>
<td>ADL/IADL (level of function)</td>
<td>The degree to which the normal physical or entire action of a system occurs (physical or psychological).</td>
</tr>
<tr>
<td>Psychosocial interaction (level of function)</td>
<td>The degree to which the normal action of a system occurs.</td>
</tr>
</tbody>
</table>

ADL/IADL, activities of daily living/instrumental activities of daily living; LPN, licensed practical nurse; LVN, licensed vocational nurse; RN, registered nurse; UAP, unlicensed assistive personnel.


The most significant of these studies, conducted by Needleman et al. (2002a, 2002b), used discharge and staffing data from 799 hospitals in 11 states to estimate nurse staffing levels for RNs, LPNs/LVNs, and aides, as well as the frequency of a wide range of complications developed by patients during their hospital stay. The data cover 6 million patients discharged from hospitals in 1997. These investigators found that low levels of RN staffing among a hospital’s nurses were associated with higher rates of serious complications such as pneumonia, upper gastrointestinal bleeding, shock, and cardiac arrest, including deaths among patients with these three complications, as well as sepsis or deep vein thrombosis. These complications occurred 3% to 9% more often than in hospitals with lower levels of RN staffing.

Sochalski (2001), after a review of studies on staff mix, cautions that
comes can only be fully achieved in the presence of other features in the practice environment, then it will be the presence or absence of these features and not solely staffing levels that will produce the desired patient results. Without a clear understanding of the circumstances under which staffing affects outcomes, we lack the capacity to improve patient outcomes if efforts are directed only at changing staffing levels. (p. 11)

Standing, Anthony, and Hertz (2001) conducted a triangulated study of outcomes after delegation to unlicensed assistive personnel (UAP), funded by the National Council of State Boards of Nursing. This report describes the qualitative analysis of interviews of RNs who described a delegation with a positive outcome and a delegation with a negative outcome. Negative outcomes after delegation ranged from family or client upsets, to fractures or other injuries, to death. In some cases, the UAP performed activities that had not been delegated to them. Negative outcomes were most frequently due to the UAP's not receiving or following directions, or not adhering to established policy. Positive outcomes included enhanced client well-being as indicated by increased socialization and other measures, prevention of poor client outcomes, and enhanced unit functioning.

Nursing-Sensitive Patient Outcomes

Donobedian states that a clear link must be established between an outcome and the process that resulted in the outcome. Thus, selecting a nursing-sensitive outcome requires a clear explanation of the process that led to that outcome. The process that needs to be defined is likely to be a complex combination of nursing acts, acts of other professionals, organizational acts, and patient characteristics and behaviors. Nursing acts are not clearly defined and are inconsistent across nurses and institutions. Few studies have attempted to describe a particular nursing process, much less link it to outcomes.

Stetler, Morsi, and Burns (2000) have worked to develop a comprehensive, in-depth profile of nursing-sensitive outcomes of hospital nursing care at the unit level and to use the information in routine quality monitoring. They used a prevention framework based on the work of Stetler and DeZell (1989). The framework describes the nurse's role in preventing complications in a nosocomial hospital environment; treatment consequences; a patient's health status, disease state, or evolving condition; and the patient's inability to care for themselves safely. From a safety perspective, the framework classifies outcomes as positive or negative. Outcomes are further classified in terms of preventability, impact, severity, and a holistic view of patient safety. Positive behaviors protect or rescue patients from potential or actual negative events. These actions are categorized as (1) detection/reporting, (2) detection/prevention, and (3) facilitation of resolution/prevention.
Other work has included development of a model of nursing effectiveness to use in studies of patient outcomes.

Irvine, Sidani, and Hall (1998) have developed The Nursing Role Effectiveness Model (Figure 9-9) to guide the examination and explanation of the links between nursing processes and patient outcomes. The model is based on Donabedian’s theory of quality care. Roles are defined as positions in organizations that have attached to them a set of expected behaviors. Professional roles are complex because they consist of components that are based on normative expectations concerning standards of practice that have been established by external regulatory bodies and secondly, on normative expectations that have evolved over time that are unique to the organization. (p. 59)

The Nursing Role Effectiveness Model has three major components: structure, the nurses’ role, and patient/health outcomes. Structure has three subcomponents: nurse, organizational, and patient. Nurse variables that influence quality of nursing care include factors such as experience level, knowledge, and skill level. Organizational components that can affect quality of nursing care include staff mix, workload, and assignment patterns. Patient characteristics that can affect quality of care include health status, severity, and morbidity. Nurses’ role has three subcomponents: nurses’ independent role, nurses’ dependent role, and nurses’ interdependent role. Independent role functions include assessment, diagnosis, nurse initiated interventions, and follow-up care. The patient/health outcomes of the independent role are clinical/symptom control, freedom from complications, functional status/self-care, knowledge of disease and its treatment, satisfaction and costs. The dependent role functions include execution of medical orders and physician-initiated treatments. It is the dependent role functions that can lead to patient/health outcomes of adverse events. Interdependent role functions include communication, case management, coordination of care, and continuity/monitoring and reporting. The interdependent role results in team functioning and affects the patient/health outcomes of the independent role.

The Propositions of The Nursing Role Effectiveness Model were stated as follows (Irvine, Sidani, & Hall, 1998):

Nursing’s capacity to engage effectively in the independent, dependent, and interdependent role functions is influenced by individual nurse variables, patient variables, and organizational structure variables. (p. 61)

Nurses’ interdependent role function depends upon the nurse’s ability to communicate and articulate her/his opinion to other members of the health care team. (p. 61)

Nurse, patient, and system structural variables have a direct effect on clinical, functional, satisfaction, and cost outcomes. (p. 61)

Nurses’ independent role function can have a direct effect on clinical, functional, satisfaction, and cost outcomes. (p. 61)

Medication errors and other adverse events associated with nurses’ dependent role function can ultimately affect all categories of patient outcome. (p. 62)

Nursing’s interdependent role function can affect the quality of interprofessional communication and coordination. The nature of inter-professional communication and coordination can influence other important patient outcomes and costs such as risk-adjusted length of stay, risk-adjusted mortality rates, excess home care costs following discharge, unplanned visits to the physician or emergency department, and unplanned re-hospitalization. (p. 62)
Sidani and Irvine (1999) subsequently modified The Nursing Role Effectiveness Model to evaluate the nurse practitioner role in acute care settings, providing a guide for nurse practitioners interested in studying their roles. These investigators selected variables relevant to acute care nurse practitioner (ACNP) practice to operationalize the components of the model, and proposed relationships among the elements of the three components. Using this framework, the research group examined the organizational factors influencing nurse practitioners’ role implementation in acute care settings (Irvine et al., 2000). They also conducted a study examining the practice patterns of ACNPs (Sidani et al., 2000). Practice patterns of ACNPs also have been examined by Rosenfeld, McEvoy, and Glassman (2003). Organizational changes resulting from implementing the APN’s role in acute care were studied by Cummings, Fraser, and Tarlier (2003).

In 2002, Doran, Sidani, Keatings, and Doidge conducted an empirical test of The Nursing Role Effectiveness Model. These investigators found that The Nursing Role Effectiveness Model was effective in guiding the evaluation of outcomes of nursing care. They noted that “for the most part the hypothesized relationships among the variables were supported. However, further work is needed to develop an understanding of how nurses engage in their coordinating role functions and how we can measure these role activities” (p. 30).

**Advanced-Practice Nursing Outcomes Research**

Advanced-practice nurses (APNs) and their organizations have been slow to move into research measuring outcomes of their practice. Early studies examined broad variables that
allow comparisons with other practitioners. Nurse-sensitive outcomes that document effectiveness of specific interventions have yet to be identified, however (Kleinpell-Nowell & Weiner, 1999). Computerized electronic patient records make the study of clinical practice considerably easier, but the use of these electronic records varies widely among APNs. APNs are very familiar with Current Procedural Terminology (CPT) codes and International Classification of Diseases-9 (ICD-9) codes. However, most are unfamiliar with standardized nursing languages, such as Nursing Outcomes Classification (NOC) and Nursing Interventions Classification (NIC), which tend to be used in electronic databases. Still, a few studies of APN practice are beginning to appear in the literature.

**RESEARCH EXAMPLE  Advanced-Practice Nursing Outcomes Research**

Barton, Gilbert, Erickson, Baramee, Sowers, and Robertson (2003) used the Omaha System, a nursing-sensitive database, to describe and document faculty practice at the University of Colorado Health Science Center. This system includes all of the nursing elements (diagnosis, intervention, outcome) and has been tested through 11 years of federally funded research. Barton et al. used Teleform software, developed by Cardiff (Vista, California) to collect data. This allowed data to be scanned, faxed, or entered using the Internet. Although the study focused on use of the system and reported only preliminary results, the project demonstrated an effective way to examine the clinical practice of APNs.

Ingersoll, McIntosh, and Williams (2000) conducted a study asking APNs what measures they recommended for use in measuring their effect on patient outcomes. Indicators listed were rated by a second group of APNs on validity, sensitivity, feasibility, utility, and cost. The highest-ranked indicators were satisfaction with care delivery, symptom resolution/reduction, perception of being well cared for, compliance/adherence with treatment plan, knowledge of patients and families, trust of care provider, collaboration among care providers, frequency and type of procedures ordered, and quality of life.

Donohue (2003) conducted a qualitative study examining nurse practitioner–client interaction in a women’s health clinic to determine what the women’s expectations were of the nurse practitioner visit and what was actually received. Clients indicated that they expected and received services, health information, trust, self-disclosure, support, affirmation, time, acceptance, and respect.

**Practice-Based Research Networks**

A practice-based research network (PBRN) is a group of practices focused on patient care that are affiliated in order to analyze their clinical practices in communities. Such networks have consisted of primary care physicians for a number of years, but no networks of advanced-practice registered nurses (APRNs) (i.e., APNs) have yet been established. Thus, questions about APRNs have not been studied. Little is known about APRNs or their practice patterns. Developing a PBRN of primary care APRNs will let researchers address questions about whether care by APRNs is different from care provided by other disciplines. In 2000, the AHRQ awarded a grant to Yale University School of Nursing in collaboration with five other schools of nursing to develop a PBRN of APRNs providing primary care. The name of the network is APRNet (Advanced Practice Registered Nurses’ Research Network). Its purpose is to “conduct and facilitate practice-based research relevant to APRN primary care practice; develop culturally competent, evidence-based practice models for APRNs; and
translate research findings into primary care practice” (McCloskey, Grey, Deshefy-Longhi, & Grey, 2003, p. 39). Initial data gathering has been primarily descriptive and examines the characteristics of practitioners participating in the network. A series of studies is planned to determine how APRN practices operate, how APRN services are determined and billed, and what clinical outcomes are obtained.

**Methodologies for Outcomes Studies**

A research tradition for the outcomes model is still emerging. A research tradition defines an acceptable research methodology. The lack of an established set of methods should encourage greater creativity in seeking new strategies for studying the phenomena of concern. Small single studies using untried methods may be useful. Research teams need to develop research programs with a planned sequence of studies focused on a particular outcome concern. The PORTs defined a research process for conducting programs of funded outcomes studies. These programs are complex and may consist of multiple studies using a variety of research strategies whose findings must be merged before conclusions are reached.

Although starting a research program as extensive as a PORT would be unrealistic without the level of funding the PORTs received, ideas for developing the methodology of outcomes research programs on a smaller scale may emerge from an examination of these plans. For example, measurement methods used in PORTs are available for smaller studies. The following steps were constructed combining PORT plans proposed by Freund, Dittus, Fitzgerald, and Heck, (1990), Sledge (1993), and Turk and Rudy (1994).

1. Perform a critical review of the published literature or a meta-analysis.
2. Conduct large database analyses on the basis of the results of the critical literature review.
3. Identify outcomes measures for use in the study, and evaluate their sensitivity to change.
4. Identify variables that might affect the outcomes.
5. Achieve consensus on definitions for all variables to be used in the research program.
6. Develop assessment instruments or techniques.
7. Conduct patient surveys or focus groups to gain information on outcomes, such as level of functional status and perceived pain, and on how these outcomes may improve or regress over time.
8. Determine patterns of care (who provides care at what points of time for what purposes?).
9. Perform a cohort analysis: Monitor a cohort of patients, some of whom will receive one treatment and others of whom will not receive the treatment, to assess changes in outcomes over time. Use a telephone survey at selected intervals to gather information. Evaluate the proportion of patients who improve, as well as the group mean differences.
10. Determine, through follow-up studies, differences in patient selection or interventions that are associated with different outcomes. Evaluate the durability of change by conducting sufficiently long follow-up. Determine the percentage of patients dropping out of groups receiving different treatments and, when possible, determine their reasons for dropping out.
11. Determine the clinical significance of improvement, as well as the statistical significance.
12. Determine the cost-benefit ratio and cost-effectiveness of the treatments under evaluation.
13. Use decision analyses to synthesize information about patients' outcomes and preferences for various types of outcomes.
14. Disseminate information to both patients and health care providers about which persons would and which would not benefit from the procedure.
15. Conduct a clinical trial to evaluate the effects of the intervention.
16. Incorporate findings into treatment guidelines.
17. Modify provider and patient behavior so that proven, effective treatment is given to persons who are most likely to benefit.

The PORTs recognized the need to allow diversity in research strategies, measures, and analyses to facilitate methodological advances (Fowler, Cleary, Magaziner, Patrick, & Benjamin, 1994). Creative flexibility often is necessary to develop ways to answer new questions. Finding ways to determine the impact of a condition on a person's life is difficult. Interpreting results also can be problematic, because clinical significance is considered as important as statistical significance. This issue requires a judgment by the research team about what constitutes clinical significance in that particular area of study.

This section describes some of the sampling issues, research strategies, measurements, and statistical approaches being used by researchers in outcomes studies. The descriptions provided are not sufficient to guide the researcher in using the approaches described but rather provide a broad overview of a variety of methods being used. For additional information, refer to the citations for each topic. Outcomes studies cross a variety of disciplines; thus, the emerging methodology is being enriched by a cross-pollination of ideas, some of which are new to nursing research.

**Samples and Sampling**

The preferred sampling methods differ in outcomes studies; random sampling is not considered desirable and is seldom used. Heterogeneous, rather than homogeneous, samples are obtained. Traditional researchers use sampling criteria that restrict the subjects to decrease possible biases and variance and to increase the possibility of identifying a statistically significant difference. Outcomes researchers, however, seek large heterogeneous samples that reflect, as much as possible, all patients who would receive care in the real world. Outcomes samples must include, for example, patients with various comorbid conditions and patients with various levels of health. In addition, persons who do not receive treatment for their condition should be identified. Evaluating the representativeness of such samples is problematic. Similarly, locating untreated patients and including them in follow-up studies constitute other challenges. Outcomes researchers must devise ways to overcome these challenges.

Traditional researchers and statisticians argue that when patients are not selected randomly, biases and confounding variables are more likely to occur. Further, they argue, this issue is a particular problem when the sample size is small. In nonexperimental studies, variation is likely to be greater, resulting in a higher risk of a Type II error. Traditional analysts consider nonrandomized studies to be based on observational data and therefore
lacking in credibility (Orchard, 1994). Using this argument, traditionalists claim that the findings of most outcomes studies are not valid and should not be used as a basis to establish guidelines for clinical practice or to build a body of knowledge.

Slade, Kuipers, and Priebe (2002) suggest that

…research questions are designed so that they can be answered by Randomized Controlled Trials (RCTs). Specifically, the use of RCTs involves the identification of an intervention which is given to patients in the experimental group, but not the control group. This encourages the asking of particular types of research questions, typically of the form “Does intervention X work for disorder Y?” However, one might argue that the RCT methodology limits the questions that can be asked, and hence can restrict the potential findings from research. Furthermore, if different questions were being asked, the RCTs would not always be the best methodology to employ…. The question “Which patients with condition Y does intervention X work for?” may prove to have more clinical relevance, and answering this question may involve asking the question “How does intervention X work?”, a question which cannot be answered just by using RCTs. (pp. 12–13)

Large Databases as Sample Sources

One source of samples used for outcomes studies is large databases. Two broad categories of databases emerge from patient care encounters: clinical databases and administrative databases, as illustrated in Figure 9-10. Providers such as hospitals, HMOs, and health care professionals create clinical databases. The clinical data are generated either as a result of routine documentation of care or in relation to a research protocol. Some databases are data registries that have been developed to gather data related to a particular disease, such as cancer (Lee & Goldman, 1989). With the use of a clinical database, it is possible to link observations made by many practitioners over long periods. Links can be made between the process of care and outcomes (Mitchell et al, 1994; Moses, 1995).

Insurance companies, government agencies, and others not directly involved in providing patient care create administrative databases. These databases have standardized sets of data for enormous numbers of patients and providers (Deyo et al., 1994; McDonald & Hui,
1991). An example is the Medicare database managed by the Health Care Financing Administration (HCFA). These large administrative databases can be used to determine the incidence or prevalence of disease, geographic variations in medical care utilization, characteristics of medical care, outcomes of care, and complementarity with clinical trials. Wray and colleagues (1995) caution, however, that analyses should be restricted to outcomes specific to a particular subgroup of patients, rather than one adverse outcome of all disease states.

Problems with the quality of data in the large databases are well recognized. The data have been gathered and entered by hundreds of people in a variety of settings. Few quality checks on the data are performed, and within the same data sets, records may have different lengths and structures. Missing data are common. Sampling and measurement errors are inherent in all large databases. Sampling error is a result of the way in which cases are selected for inclusion in the database; measurement error emerges from problems related to the operational definition of concepts. Thus, reliability and validity of the data are concerns (Davis, 1990; Lange & Jacox, 1993).

Large databases are used in outcomes studies to examine patient care outcomes. The outcomes that can be examined are limited to those recorded in the database and thus tend to be rather general. Existing databases can be used for analyses such as assessment of nursing care delivery models, variation in nursing practices, or evaluation of patients’ risk of hospital-acquired infection, hospital-acquired pressure ulcer, or falls. Lange and Jacox (1993) identify the following important health policy questions related to nursing that should be examined through the use of large databases:

1. What is standard nursing practice in various settings?
2. What is the relationship between variations in nursing practice and patient outcomes?
3. What are the effects of different nursing staff mixes on patient outcomes and costs?
4. What are the total costs for episodes of treatment of specific conditions, and what part of those are attributable to nursing care?
5. Who is being reimbursed for nursing care delivery? (Lange & Jacox, p. 207)

To examine these questions, nurses must develop the statistical and methodological skills needed for working with large databases. Large databases contain patient and institutional information from huge numbers of patients. They exist in computer-readable form, require special statistical methods and computer techniques, and can be used by researchers who were not involved in the creation of the database.

Regrettably, nursing data are noticeably missing from these large databases and hence from funded health policy studies using them. A nursing minimum data set has been repeatedly recommended for inclusion in these databases (Werley, Devine, Zorn, Ryan, & Westra, 1991; Werley & Lang, 1988; Zielstorff, Hudgings, Grobe, & the National Commission on Nursing Implementation Project Task Force on Nursing Information Systems, 1993). This minimum data set would comprise a set of variables necessary and sufficient to describe an episode of illness or the care given by a provider. The ANA has mandated the formation of a Steering Committee on Databases to Support Clinical Nursing Practice. The following nursing classification schemes are being used in national databases:

- The North American Nursing Diagnosis Association (NANDA) classification
- The Omaha System: Applications for Community Health Nursing classification
- The Home Health Care Classification
- The Nursing Interventions Classification (NIC)
The Nursing Outcomes Classification (NOC)

Temple (1990) expressed the following concerns regarding the use of large data sets, rather than controlled trials, to assess effectiveness of treatments:

We have traveled this route before with uncontrolled observations. It has always been hoped, and has often been asserted, that uncontrolled databases can be adjusted in some way that will allow valid comparisons of treatments. I know of no systematic attempt to document this. Outcomes researchers counter these criticisms by pointing out that experimental studies lack external validity and are not useful for application in clinical settings. They claim that the findings of clinical trials are not being used by clinicians because they are not representative of the patients seeking care. (p. 211)

Research Strategies for Outcomes Studies

Outcomes research programs usually consist of studies with a mix of strategies carried out sequentially. Although these strategies could be referred to as designs, for some the term design as used in Chapter 8 is inconsistent with the use of the term here. Research strategies for outcomes studies have emerged from a variety of disciplines, and innovative new strategies continue to appear in the literature. Strategies for outcomes studies tend to allow less control than is possible with traditional research designs and cannot be as easily categorized. The numerous research strategies for outcomes studies described next are only a sampling from the literature:

- Consensus knowledge building
- Practice pattern profiling
- Prospective cohort studies
- Retrospective cohort studies
- Population-based studies
- Clinical decision analysis
- Study of the effectiveness of interdisciplinary teams
- Geographical analyses
- Economic studies
- Ethical studies
- Defining and testing of interventions

Consensus Knowledge Building

Consensus knowledge building usually is performed by a multidisciplinary group representing a variety of constituencies. Initially, an extensive international search of the literature on the topic of concern, including unpublished studies, studies in progress, dissertations, and theses, is conducted. Several separate reviews may be performed, focusing on specific questions about the outcomes of care, diagnosis, prevention, or prognosis. Because meta-analytic methods often cannot be applied to the literature pertinent to PORTs, systematic approaches to critique and synthesis have been developed to identify relevant studies and gather and analyze data abstracted from the studies (Powe et al., 1994).

The results are dispersed to researchers and clinical experts in the field, who are asked to carefully examine the material and then participate in a consensus conference. The consensus
conference yields clinical guidelines, which are published and widely distributed to clinicians. The clinical guidelines also are used as practice norms to study process and outcomes in that field. Gaps in the knowledge base are identified and research priorities determined by the consensus group.

Preliminary steps in this process may include conducting extensive integrative reviews and seeking consensus from a multidisciplinary research team and locally available clinicians. A review can be accomplished by establishing a website and conducting dialogue with experts via the Internet. The review may then be published in Sigma Theta Tau’s online journal, *Knowledge Synthesis in Nursing*, and then dialogue related to the review may be conducted over the Internet. The Delphi method also has been used to seek consensus (Vermeulen, Ratko, Erstad, Brecher, & Matuszewski, 1995).

**Practice Pattern Profiling**

*Practice pattern profiling* is an epidemiological technique that focuses on patterns of care, rather than on individual occurrences of care. Large database analysis is used to identify a provider’s pattern of practice and compare it with that of similar providers or with an accepted standard of practice. The technique has been used to determine overutilization and underutilization of services, to determine costs associated with a particular provider’s care, to uncover problems related to efficiency and quality of care, and to assess provider performance. The provider being profiled may be an individual practitioner, a group of practitioners, or a health care organization such as a hospital or an HMO.

The provider’s pattern is expressed as a rate aggregated over time for a defined population of patients under the provider’s care. For example, the analysis may examine the number of sigmoidoscopy claims filed per 100 Medicare patients seen by the provider in a given year. Other analyses may examine (1) whether diabetic patients have had at least one annual serum glucose test and have received an ophthalmology examination or (2) the frequency of flu shots, Papanicolaou smears, and mammograms for various target populations (Lasker, Shapiro, & Tucker, 1992; McNeil, Pedersen, & Gatsonis, 1992).

Profiling can be used when the data contain hierarchical groupings. For example, patients may be grouped by nurse, nurses by unit, and units by larger organizations. The analysis uses regression equations to examine the relationship of an outcome to the characteristics of the various groupings. To be effective, the analysis must include data on the different sources of variability that might contribute to a given outcome.

The structure of the analysis reflects the structure of the data. Patient characteristics, for example, may include data on disease severity, comorbidity, emergent status, behavioral characteristics, socioeconomic status, and demographics. Nurse characteristics may consist of level of education, specialty status, years of practice, age, gender, and certifications. Unit characteristics may comprise number of beds, nursing management style used on the unit, ratio of patients to nurses, and the proportion of staff who are registered nurses (RNs) (McNeil et al., 1992).

Profiles are designed to generate some type of action, such as demonstrating that a provider’s rates are too high or too low compared with the norm. By examining aggregate patterns of practice, profiling can be used to evaluate the care provided by different organizations or received by different populations of patients. Critical pathways or care maps can then be used to determine the proportion of patients whose data diverged from the pathway
for a particular nurse, group of nurses, or group of nursing units. Profiling can be used for quality improvement, assessment of provider performance, and utilization review.

Methods of improving outcomes are not addressed by profiling, although this process can identify problem areas. It can be used to determine how and by whom performance should be changed to improve outcomes. Profiling also can identify outliers (persons with extreme scores or values), allowing more detailed examination of the reasons for the disparity of data.

The databases currently being used for profiling are not ideal, because they were developed for other purposes. Outcomes that can be examined are limited to broad outcomes, such as morbidity and mortality rates, complication rate, readmission rate, and frequency of utilization of various services (Lasker et al., 1992; McNeil et al., 1992). Table 9-4 lists examples of the large database measures that might be used in profiling.

<table>
<thead>
<tr>
<th>Quality of Care Issue</th>
<th>Measures</th>
<th>Example(s)</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access</td>
<td>Proportion of population receiving care during the year, classified by age and sex</td>
<td>% of children under age 2 seen for at least one well-care visit</td>
<td>National Trends</td>
</tr>
<tr>
<td></td>
<td></td>
<td>% of children seen in emergency rooms for any reason, for trauma, and for medical problems</td>
<td></td>
</tr>
<tr>
<td>Preventive</td>
<td>Portion of population in specific age and sex groups receiving recommended tests or procedures</td>
<td>% of children by group having recommended immunizations in previous year</td>
<td>National recommendation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>% of women age 50 and over having mammography in past year</td>
<td>National recommendation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>% of deliveries with prenatal care beginning in first trimester</td>
<td>National recommendation</td>
</tr>
<tr>
<td>Diagnosis</td>
<td>% of population diagnosed (and under care) for specific chronic conditions by age and sex</td>
<td>% of adults diagnosed at one or more visits as having essential hypertension by age and sex</td>
<td>Epidemiologic data on prevalence of hypertension</td>
</tr>
<tr>
<td>Treatment</td>
<td>Medications</td>
<td>Average number of new prescriptions per person per year</td>
<td>Average number of new prescriptions for antibiotics per person per year</td>
</tr>
<tr>
<td></td>
<td>Surgery</td>
<td>Rate of surgical procedures per year: total, inpatient, and ambulatory (if applicable)</td>
<td>Cesarean section rate for all deliveries</td>
</tr>
<tr>
<td>Outcomes</td>
<td>Hospital readmissions within 3 months of discharge</td>
<td>% of readmissions for some condition</td>
<td>Comparison data and trends</td>
</tr>
<tr>
<td></td>
<td></td>
<td>% of readmission identifying a complication</td>
<td></td>
</tr>
</tbody>
</table>

Prospective Cohort Studies

A prospective cohort study is an epidemiological study in which a group of people are identified who are at risk for experiencing a particular event. Sample sizes for these studies often must be very large, particularly if only a small portion of the at-risk group will experience the event. The entire group is followed over time to determine the point at which the event occurs, variables associated with the event, and outcomes for persons who experienced the event compared with those who did not.

The Harvard Nurses Health Study, which is still being conducted, is an example of a prospective cohort study. This study recruited 100,000 nurses to determine the long-term consequences of the use of birth control pills. Nurses are sent a questionnaire every 2 years to gather data about their health and health behaviors. The study has been in progress for more than 20 years. Multiple studies using the large data set yielded by this Harvard study have been reported in the literature. Prospective cohort nursing studies could be conducted on a smaller scale on other populations, such as patients identified as being at high risk for the development of pressure ulcers.

Retrospective Cohort Studies

A retrospective cohort study is an epidemiological study in which a group of people are identified who have experienced a particular event. This is a common research technique used to study occupational exposure to chemicals. Events of interest to nursing that can be studied in this manner include a procedure, an episode of care, a nursing intervention, and a diagnosis. For example, nurses may use a retrospective cohort study to follow a cohort of women who have received a mastectomy for breast cancer or of patients in whom a urinary bladder catheter was placed during and after surgery. The cohort is evaluated after the event to determine the occurrence of changes in health status, usually identified as the development of a particular disease or death. Nurses also may be interested in the pattern of recovery after an event or, in the case of catheterization, the incidence of bladder infections in the months after surgery.

On the basis of the study findings, epidemiologists calculate the relative risk of the identified change in health for the group. For example, if death is the occurrence of interest, the expected number of deaths is determined. The observed number of deaths divided by the expected number of deaths and multiplied by 100 yields a standardized mortality ratio (SMR), which is regarded as a measure of the relative risk for the persons in the studied group to die of a particular condition. For example, in nursing studies, patients may be evaluated at specific intervals after discharge from a health care facility (Swain & Meijers, 1988).

In retrospective studies, researchers commonly ask patients to recall information relevant to their previous health status. This information often is used to determine the amount of change occurring before and after an intervention. However, because recall can easily be distorted, and researchers thereby misled, retrospective studies should be used with caution. Herrmann (1995) identified three sources of distortion in recall, as follows: (1) the question posed to the subject may be conceived or expressed incorrectly, (2) the recall process may be in error, and (3) the research design used to measure recall can result in the recalled event's appearing to be different from what actually occurred. Herrmann (1995, p. AS90) also identified four bases of recall:
Direct recall: the subject “accesses the memory without having to think or search memory” resulting in correct information.

Indirect recall: the subject “accesses the memory after thinking or searching memory,” resulting in correct information.

Limited recall: “access to the memory does not occur but information that suggests the contents of the memory is accessed,” resulting in an educated guess.

No recall: “neither the memory nor information relevant to the memory may be accessed,” resulting in a wild guess.

Population-Based Studies

Population-based studies also are important in outcomes research. Conditions must be studied in the context of the community, rather than of the medical system. To avoid selection bias with this method, all cases of a condition occurring in the defined population are included, rather than only patients treated at a particular health care facility. Efforts may be made to include people with the condition who have not received treatment.

Community-based norms of tests and survey instruments obtained in this manner provide a clearer picture of the range of values than does evaluation of the limited spectrum of patients seen in specialty clinics. Estimates of instrument sensitivity and specificity are more accurate. This method is useful in elucidating the natural history of a condition or in identifying the long-term risks and benefits of a particular intervention (Guess, Jacobsen, Girman, Oesterling, Chute, Panser, et al., 1995).

Clinical Decision Analysis

Clinical decision analysis is a systematic method of describing clinical problems, identifying possible diagnostic and management courses of action, assessing the probability and value of various outcomes, and then calculating the optimal course of action. Decision analysis is based on the following four assumptions: (1) decisions can be quantified; (2) all possible courses of action can be identified and evaluated; (3) the different values of outcomes, viewed from the perspective of the nurse, patient, payer, and administrator, can be examined; and (4) the analysis allows selection of an optimal course of therapy.

To perform the analysis, the researchers must define the boundaries of the clinical in terms of a logical sequence of events over time. All possible courses of action are then determined. These courses of action usually are represented in a decision tree consisting of a starting point, available alternatives, probable events, and outcomes. Next, the goals and objectives of problem resolution are defined. Researchers calculate the probability of occurrence of each path of the decision tree and ensure that an outcome exists for each potential path. Each outcome is assigned a value. These values may be expressed in terms of money, morbidity incidents, quality-of-life measures, or duration of hospital stay. (Figure 9-11 displays a simplified decision tree for breech delivery in obstetrics.) Researchers can then identify an optimal course of action according to which decision maximizes the chances of the most desirable outcomes (Crane, 1988; Keeler, 1994; Sonnenberg, Roberts, Tsevat, Wong, Barry, & Kent, 1994).

Studies analyzing clinical decisions have primarily used questionnaires and interviews. However, determining the clinical decisions of practitioners is not an easy task. Much of
patient care involves the clinician and the patient alone. The underlying theories of care and the processes of care are hidden from view. Thus, it is difficult for clinicians to compare their specific approaches to care. Among physicians, care delivered by other physicians is rarely observed (O’Connor et al., 1993). Studies have found that physicians have difficulty recalling their decisions and providing rationales for them (Chaput de Saintonge & Hattersley, 1985; Kirwan, Chaput de Saintonge, Joyce, Holmes, & Currey, 1986).

Unsworth (2001) describes the following strategies for studying clinical decision making:

- semistructured interviews
- audio-assisted recall (the clinician listens to an audiotape of the clinician-client dialogue and uses this to aid recall of her or his reasoning processes)
- video-assisted recall (the clinician uses video footage to prompt recall of reasoning processes)
- the clinician writing notes as he or she solves a problem
- the think-aloud method (the clinician provides verbal commentary during interaction with the client)
- the clinician presents the reasoning about a clinician-client session afterward from memory use of a head-mounted video camera with video-assisted recall.

Chaput de Saintonge and associates (1988) propose a strategy for analyzing clinical decisions using “paper patients.” The techniques seem to parallel the decisions made by
practitioners in the clinical setting. In their study, 10 common clinical variables used to evaluate the status of patients with rheumatoid arthritis were collected at two points for 30 patients participating in a clinical trial, at the time of entry and 1 year later. Data for 20 of the patients were duplicated throughout the table to check the consistency of responses, making 50 responses in all. The variables were presented to rheumatologists on a single sheet of paper labeled “before” and “after a year.” Physicians were asked to indicate the extent of change in each patient’s condition using a visual analogue scale (VAS) with the ends labeled “greatest possible deterioration” and “greatest possible improvement.” They also were asked whether they considered the change clinically important. Then they were asked to indicate the relative importance of each variable, rating the variables on a scale of 1 to 100.

Regression analyses were performed in which each VAS variable was used as the dependent variable. With increasing VAS values, judgments of clinical importance changed from “not important” to “important.” This change occurred over a 5-mm length of the scale or less. (VAS scales traditionally are 100 mm in length.) The researchers designated the midpoint of this transition zone as the “threshold value of clinical importance.” Consistency of responses was tested to correlate responses with those of duplicate cases. The researchers then developed a consensus model by weighing each physician's responses on the basis of the correlation. The VAS scores were multiplied by the correlation coefficient. These VAS scores were then used as the dependent variable in another regression analysis. This method is useful in identifying the variables important in the making of clinical decisions and the consistency with which practitioners make their decisions.

**Study of the Effectiveness of Interdisciplinary Teams**

According to Schmitt, Farrell, and Heinemann (1988), *interdisciplinary teams* have the following characteristics:

...(1) multiple health disciplines are involved in the care of the same patients, (2) the disciplines encompass a diversity of dissimilar knowledge and skills required by the patients, (3) the plan of care reflects an integrated set of goals shared by the providers of care, and (4) the team members share information and coordinate their services through a systematic communication process. (p. 753)

Part of the communication process consists of regularly scheduled face-to-face meetings. The assumption is that collaborative team approaches provide more effective care than that delivered using team approaches or noncollaborative multidisciplinary approaches (parallel care).

Interdisciplinary teams are becoming more common as health care changes. Examples are hospice care teams, home health teams, and psychiatric care teams. Studying the effectiveness of interdisciplinary teams is difficult, however. The characteristics that make team care more effective have not been identified. Studies usually focus on the evaluation of a single team, rather than on conducting comparison studies. The outcomes of team care also are multidimensional, requiring the use of multiple dependent variables.

Evaluation studies examining team care often examine only posttreatment data without baseline data. Comparison of groups will not reveal any evidence that the groups were similar in terms of important variables before the intervention. Involvement of family
members with the team has not been examined. Clearly, this is an important focus of research requiring more rigorous designs than have previously been used.

**Geographical Analyses**

**Geographical analyses** are used to examine variations in health status, health services, patterns of care, or patterns of use by geographical area and sometimes are referred to as **small area analyses**. Variations may be associated with sociodemographic, economic, medical, cultural, or behavioral characteristics. Locality-specific factors of a health care system, such as capacity, access, and convenience, may play a role in explaining variations. The social setting, environment, living conditions, and community also may be important factors.

The interactions between the characteristics of a locality and of its inhabitants are complex. The characteristics of the total community may transcend the characteristics of persons within the community and may influence subgroup behavior. High education levels in the community commonly are associated with greater access to information and receptiveness to ideas from outside the community.

Regression analyses commonly are used to develop models using all of the risk factors and the characteristics of the community. Results often are displayed through the use of maps (Kieffer, Alexander, & Mor, 1992). After the analysis, the researcher must determine whether differences in rates are due to chance alone and whether high rates are too high. From a more theoretical perspective, the researcher must then explain the geographical variation uncovered by the analysis (Volinn, Diehr, Ciol, & Loeser, 1994).

Geographical information systems (GISs) can provide an important tool for performing geographic analyses. A GIS uses relational databases to facilitate processing of spatial information. The software tools in a GIS can be used for mapping, data summaries, and analysis of spatial relationships. GISs have the capability of modeling data flows so that the effect of proposed changes in interventions applied to individuals or communities on outcomes can be modeled (Auffrey, 1998).

**Economic Studies**

Many of the problems studied in health services research address concerns related to the efficient use of scarce resources and, thus, to economics. Health economists are concerned with the costs and benefits of alternative treatments or ways of identifying the most efficient means of care. The economist’s definition of **efficiency** is the least-cost method of achieving a desired end with the maximum benefit to be obtained from available resources. If available resources must be shared with other programs or other types of patients, an economic study can determine whether changing the distribution of resources will increase total benefit or welfare.

To determine the efficiency of a treatment, the economist conducts a **cost-effectiveness analysis**. This technique uses a single measure of outcomes, and all other factors are expressed in monetary terms as net cost per unit of output (Ludbrook, 1990). Cost-effectiveness analyses compare different ways of accomplishing a clinical goal, such as diagnosing a condition, treating an illness, or providing a service. The alternative approaches are compared in terms of costs and benefits. The purpose is to identify the strategy that provides the most value for the money. Tradeoffs between costs and benefits are unavoidable, however (Oster,

It is time for nurses to take a more active role in conducting cost-effectiveness research. Nurses are well positioned to evaluate health care practices and have the incentive to conduct the studies. Nursing practice is seldom a subject of cost-effectiveness analyses. Nevertheless, such knowledge would enable nurses to refine their practice, substituting interventions that maximize nurses’ time to the best advantage, in terms of the patient’s health, for interventions that offer less gain (Siegel, 1998).

As Lieu and Newman (1998) point out:

“[C]ost effective” does not necessarily mean “cost-saving” (Doubilet, Weinstein, & McNeil, 1986). Many health interventions, even preventive ones, do not save money (Tengs, Adams, Pliskin, Safran, Siegel, Weinstein, et al., 1995). Rather, a service should be called cost-effective if its benefits are judged worth the costs. Recently, a consensus panel supported by the National Institutes of Health published recommendations that define standards for conducting cost-effectiveness analysis (Gold, Siegel, Russell, & Weinstein, 1996). Cost-effectiveness analysis is only one of several methods that can be used for the economic evaluation of health services (Drummond, Stoddart, and Torrance, 1987). Although these methods are useful, an intervention cannot be cost-effective without being effective” (Lieu & Newman, 1998, p. 1043).

To examine overall benefits, a **cost-benefit analysis** is performed. With this method, the costs and benefits of alternative ways of using resources are assessed in monetary terms, and the use that produces the greatest net benefit is chosen. The costs included in an economic study are defined in exact ways. The actual costs associated with an activity, not prices, must be used. Cost is not the same as price. In most cases, price is greater than cost. **Costs** are a measure of the actual use of resources, rather than the price charged. Charges are a poor reflection of actual costs. Costs typically included in a cost-benefit analysis are costs to the provider, costs to third-party payers (e.g., insurance), out-of-pocket costs, and opportunity costs.

**Out-of-pocket costs** are those expenses incurred by the patient or family members, or both, that are not reimbursable by the insurance company. Examples are costs of buying supplies, dressings, medications, and special food, transportation expenses, and unreimbursable care expenses.

**Opportunity costs** are lost opportunities that the patient, family member, or others experience. For example, a family member who must stay at home to care for the patient may lose the opportunity to earn more money. A teenager who needs to drop out of school for a semester to care for a parent may lose the opportunity to advance her education. A husband might have been able to take a better job if the family could have moved to another town, rather than staying in place to enable a member to receive specific medical care.

Opportunity costs often are not included in the consideration of overall costs. This omission results in an underestimation of costs and an overestimation of benefit. For example, caring for an acutely ill patient at home is cost-effective if out-of-pocket costs and opportunity costs are not considered. However, the total costs of providing the care, regardless of who pays or who receives the money, must be included. In performing such a study, it is important to state whose costs are being considered and who is to weigh the benefits against the consequences.

Allred, Arford, Mauldin, and Goodwin (1998) critiqued the seven nursing studies between 1992 and 1996 in which cost-effectiveness analyses were performed. They found
these studies to be equivalent in quality to those from other disciplines. They concluded that more emphasis must be placed on cost-effectiveness analyses in nursing research, and they provided guidelines for conducting these studies.

Stone (1998) has described the recommended guidelines for journal reports of cost-effectiveness analyses. Cost-effectiveness studies should be used as aids in decision making, rather than as the end decision. If a cost-effectiveness study is conducted to inform those who make resource allocation decisions, a standard reference case should be presented to allow the decision makers to compare a proposed new health intervention with existing practice.

**Ethical Studies**

Outcomes studies often result in the development of policies for the allocation of scarce resources. Ethicists take the position that moral principles, such as justice, must be considered as constraints on the use of costs and benefits, to choose treatments that maximize the benefit per unit cost. Value commitments are inherent in choices about research methods and about the selection and interpretation of outcome variables, and these commitments should be acknowledged by researchers.

The choices researchers make should be documented and the reasons for those choices should be given explicitly in publications and presentations so that readers and other users of the information are enabled and expected to bear more responsibility for interpreting and applying the findings appropriately. (Lynn & Virnig, 1995)

Veatch (1993) proposes that analysis of the implications of rationing decisions in terms of the principles of justice and autonomy will provide more acceptable criteria than outcomes predictors alone. As an example, Veatch performs an ethical analysis of the use of outcome predictors in decisions related to early withdrawal of life support. Ethical studies should play an important role in outcomes programs of research.

**Measurement Methods**

The selection of appropriate outcome variables is critical to the success of a study (Bernstein & Hilborne, 1993). As in any study, evidence of validity and reliability of the methods of measurement must be evaluated. Outcomes selected for nursing studies should be those most consistent with nursing practice and theory (Harris & Warren, 1995). In some studies, rather than selecting the final outcome of care, which may not occur for months or years, measures of intermediate end points are used. **Intermediate end points** are events or markers that act as precursors to the final outcome. It is important, however, to document the validity of the intermediate end point in predicting the outcome (Freedman & Schatzkin, 1992). In early outcomes studies, researchers selected outcome measures that could be easily obtained, rather than those most desirable for outcomes studies.

Table 9-5 identifies characteristics important to evaluate in selecting methods of measuring outcomes. In evaluating a particular outcome measure, the researcher should consult the literature for previous studies that have used that particular method of measurement, including the publication describing development of the method of measurement. Information related to the measurement can be organized into a table such as Table 9-6, allowing easy comparison of several methods of measuring a particular outcome.
### Table 9-5 Characteristics of Outcomes Assessment Instruments

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Considerations in Patient Outcomes Evaluation</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Applicability</strong></td>
<td>Consider purpose of instruments</td>
<td>Deyo &amp; Carter, 1992</td>
</tr>
<tr>
<td></td>
<td>Discriminate between subjects at a point in time</td>
<td>Stewart et al., 1989</td>
</tr>
<tr>
<td></td>
<td>Predict future outcomes</td>
<td>Guyatt, Walter, &amp; Norman, 1987</td>
</tr>
<tr>
<td></td>
<td>Evaluate changes within subjects over time</td>
<td>Feinstein, Josephy, &amp; Wells, 1986</td>
</tr>
<tr>
<td></td>
<td>Screen for problems</td>
<td>Deyo, 1984</td>
</tr>
<tr>
<td></td>
<td>Assess quality of care</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Consider whether:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Norms are established for clinical population of interest</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Instrument format is compatible with assessment approach (e.g., observer rated vs. self-administered)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Setting in which instrument was developed</td>
<td></td>
</tr>
<tr>
<td><strong>Practicality (clinical utility)</strong></td>
<td>The instrument:</td>
<td>Leidy, 1991</td>
</tr>
<tr>
<td></td>
<td>Includes outcomes important to the patient</td>
<td>Nelson, Landgraf, Hays, Wasson, &amp; Kirk, 1990</td>
</tr>
<tr>
<td></td>
<td>Is short and easy to administer (low respondent burden)</td>
<td>Stewart et al., 1989</td>
</tr>
<tr>
<td></td>
<td>Questions are easy to understand and acceptable to patients and interviewers</td>
<td>Lohr, 1988</td>
</tr>
<tr>
<td></td>
<td>Scores reflect condition severity, condition-specific features, and discriminate those with conditions from those without</td>
<td>Bombardier &amp; Tugwell, 1987</td>
</tr>
<tr>
<td></td>
<td>Is easily scored and scores are readily understandable</td>
<td>Feinstein et al., 1986</td>
</tr>
<tr>
<td></td>
<td>Level of measurement allows a change score to be determined</td>
<td>Kirshner &amp; Guyatt, 1985</td>
</tr>
<tr>
<td></td>
<td>Provides information that is clinically useful</td>
<td>Deyo, 1984</td>
</tr>
<tr>
<td></td>
<td>Includes patient rating of magnitude of effort and support needed for performance of physical tasks</td>
<td></td>
</tr>
<tr>
<td><strong>Comprehensiveness</strong></td>
<td>Generic measures are designed to summarize a spectrum of concepts applied to different impairments, illnesses, patients, and populations</td>
<td>Nelson et al., 1990</td>
</tr>
<tr>
<td></td>
<td>Disease-specific measures are designed to assess specific patients with specific conditions or diagnoses</td>
<td>Patrick &amp; Deyo, 1989</td>
</tr>
<tr>
<td></td>
<td>Dimensions of the instrument; a core set of physical, mental, and role function desirable</td>
<td>Deyo, 1984</td>
</tr>
<tr>
<td><strong>Reliability</strong></td>
<td>Can be influenced by day-to-day variations in patients, differences between observers, items in the scale, mode of administration</td>
<td>Nelson et al., 1990</td>
</tr>
<tr>
<td></td>
<td>This is the critical determinant of usefulness of an instrument</td>
<td>Spitzer, 1987</td>
</tr>
<tr>
<td></td>
<td>Designed for discriminative purpose</td>
<td>Guyatt et al., 1987</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Deyo, 1984</td>
</tr>
</tbody>
</table>
Outcomes researchers are moving away from classical measurement theory as a means of evaluating the reliability of measurement methods. They are interested in identifying change in measures over time in a subject, and instruments developed through the use of classical measurement theory are often not sensitive to these changes. The magnitude of change that can be detected also is important to determine. In addition, measures may detect change within a particular range of values but may not be sensitive to changes outside that range. The sensitivity to change of many commonly used outcome measures has not been examined (Deyo & Carter, 1992; Felson, Anderson, & Meenan, 1990). Studies must be conducted specifically to determine the sensitivity of measures before they are used in outcomes studies. As the sensitivity of a measure increases, statistical power increases, allowing smaller sample sizes to detect significant differences.

Creative methods of collecting data on instruments for large outcomes studies must be explored. In a busy office or clinic setting, the typical strategy of having clerks or other staff administer questionnaires or scales to patients is time intensive and costly and may result in lost data. Greist and colleagues (1997) recommend using the computer and the telephone to collect such data. Computers containing the instrument can be placed in locations convenient to patients, so the instrument can be completed with a minimum of staff involvement.

Another option is telephone interviews using the computer. The traditional telephone interview using interviewers to ask questions is costly. The same interactive voice response (IVR) technology used in voice mail, however, can be used in telephone interviewing by computer. IVR allows the patient to respond to yes–no and multiple-choice questions by pressing numbers on the keypad or by saying “yes” or “no” or a number from 0 to 9. Patients can record answers in their own voice.

Table 9-5  Characteristics of Outcomes Assessment Instruments—cont’d

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Considerations in Patient Outcomes Evaluation</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Validity</td>
<td>No consensus of what are scientifically admissible criteria for many indices No “goal standard” exists for establishing criterion validity for many indices</td>
<td>Spitzer, 1987  Deyo, 1984</td>
</tr>
<tr>
<td>Responsiveness</td>
<td>Not yet indexed for virtually any evaluative measures Coarse scale rating may not detect changes Aggregated scores may obscure changes in subscales Useful for determining sample size and statistical power Reliable instruments are likely to be responsive but reliability not adequate as sole index of consistent results over time Consider detail in scaling As baseline variability of score changes within stable subjects, may need larger treatment effects to demonstrate efficacy Consider temporal relationship between intervention and outcome</td>
<td>Stewart &amp; Archbold, 1992  Leidy, 1991  Jaeschke, Singer, &amp; Guyatt, 1989  Guyatt et al., 1989  Bombardier &amp; Tugwell, 1987  Guyatt et al., 1987  Deyo &amp; Centor, 1986  Deyo, 1984</td>
</tr>
</tbody>
</table>

Measuring the frequency and nature of care activities of various staff has been problematic in studies of the process of care. Strategies commonly used are chart review, time and motion studies, work sampling, and retrospective recall. None of these is a satisfactory indicator of the actual care that occurs (Hale, Thomas, Bond, & Todd, 1997). Holmes, Teresi, Lindeman, and Glandon (1997) recommend the use of barcode methodology to measure service inputs. Scanning the barcodes captures what care is provided, for whom, by whom, and at what time. Barcoded service sheets and a portable barcode reader are used with an accompanying database management system.

### The Analysis of Measurement Reliability

Estimating the reliability of outcome measures through the use of classical measurement theory may be problematic. The traditional concept of measurement reliability was developed to evaluate quantities that were not expected to change over time for an individual subject. This assessment of reliability is irrelevant or only partially relevant to assessing the suitability or precision of measures selected because of their sensitivity to change within the subject over time. Traditional evaluations of measurement methods assume that any change in group

---

**Table 9–6** Characteristics of the Katz Activities of Daily Living (ADL) Scale: A Proposed Outcome Instrument

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Applicability</strong></td>
<td>Katz et al., 1970</td>
</tr>
<tr>
<td>Purpose is to objectively evaluate results of treatment in chronically ill and aging populations</td>
<td>Winner et al., 1990</td>
</tr>
<tr>
<td>Predicts service utilization in elderly population</td>
<td>Fries, 1990</td>
</tr>
<tr>
<td>Used in case-mix adjustments</td>
<td>Spector, 1990</td>
</tr>
<tr>
<td>Scale discriminates well on disability in elderly population, norms easily referenced</td>
<td>Spector, 1990</td>
</tr>
<tr>
<td>Ratings judgment based on direct observation and caregiver reports, known differences in observed vs. reported ratings</td>
<td>Burns, 1992</td>
</tr>
<tr>
<td><strong>Practicality</strong></td>
<td>Katz et al., 1970</td>
</tr>
<tr>
<td>Brief, 6 items with 3 levels of dependency</td>
<td>Spector, 1990</td>
</tr>
<tr>
<td>Can be used by clinicians and non-clinicians</td>
<td>Katz et al., 1970</td>
</tr>
<tr>
<td>Measures performance (not ability)</td>
<td>Spector, 1990</td>
</tr>
<tr>
<td>Aggregate score represents increasing level of dependency</td>
<td>Burns, 1992</td>
</tr>
<tr>
<td><strong>Comprehensiveness</strong></td>
<td>Katz et al., 1970</td>
</tr>
<tr>
<td>Includes bathing, dressing, toileting, transfer, continence, and eating</td>
<td>Kane &amp; Bayer, 1991</td>
</tr>
<tr>
<td>Does not explain etiology of level of performance</td>
<td>Kane &amp; Bayer, 1991</td>
</tr>
<tr>
<td><strong>Reliability</strong></td>
<td>Kane &amp; Bayer, 1991</td>
</tr>
<tr>
<td>Performance may be influenced by motivational, social, and environmental factors</td>
<td>Spector, 1990</td>
</tr>
<tr>
<td>High internal consistency reported</td>
<td>Spector, 1990</td>
</tr>
<tr>
<td><strong>Validity</strong></td>
<td>Spector, 1990</td>
</tr>
<tr>
<td>Content and construct validity assessments are acceptable</td>
<td>Spector, 1990</td>
</tr>
<tr>
<td><strong>Responsiveness</strong></td>
<td>Spector, 1990</td>
</tr>
<tr>
<td>No published reports that quantify relationship of scale change to minimal clinically important change</td>
<td></td>
</tr>
</tbody>
</table>

values is a result of interindividual variation. Patient change, however, results in changes within the subject. With classical measurement theory analysis, a measure that did not vary between individual subjects is considered to have zero (or poor) reliability. This measure, however, may be an excellent measure of change over time if individual subjects change on that measure (even if group averages do not change much). Thus, it is inappropriate to assess the reliability of difference scores according to the internal consistency of measures (Collins & Johnston, 1995).

Some outcomes researchers use measures obtained from individual subjects as indicators of characteristics of a group. The data from the measures are aggregated to reflect the group. In this case, the researcher must assess the extent to which the responses represent the group. Although the group mean usually is expected to serve this purpose, it may not adequately represent the group. Verran, Mark, and Lamb (1992) describe techniques for examining the psychometric properties of instruments used to describe group-level phenomena. Items of the instrument should be assessed for content validity to determine how well they measure group-level concepts. Reliability and validity must be assessed at the aggregated level, rather than at the individual level.

Commonly, multiple outcomes measures are used in outcomes studies. Researchers wish to evaluate all relevant effects of care. However, quantity of measures is not necessarily evidence of the quality of the measures. Researchers should select the measures most relevant to the treatment, avoiding measures that are closely correlated. Interpreting the results of studies in which multiple outcomes have been used can be problematic. For example, Felson and colleagues (1990, p. 141) ask “which is the better therapy, the one that shows a change in 6 outcome measures out of 12 tested or the one that shows a change in 4 of the 12 measures? What if the 4 that demonstrate change with one therapy are not the same as the 6 that show a change in another therapy?” If multiple comparisons are made, it is important to make statistical adjustments for them; the risk of a Type I error is greater when multiple comparisons are made.

Some researchers recommend combining various measures into a single summary score (DesHarnais, McMahon, & Wroblewski, 1991; Felson et al., 1990). Such global composite measures have not been widely used, however. The various measures used in such an index may not be equally weighted and may be difficult to combine. Also, the composite index value may not be readily interpretable by clinicians.

The focus of most measures developed for outcomes studies has been the individual patient. However, a number of organizations are now developing measures of the quality of performance of systems of care. In 1990, the Consortium Research on Indicators of System Performance (CRISP) project began to develop indicators of the quality of performance of integrated delivery systems. From the perspective of CRISP, the success of a health system is associated with its ability to decrease the number of episodes of diseases in the population. Therefore, the impact of the delivery system on the community is considered an important measure of performance. CRISP has developed a number of indicators now in use by consortium members, who pay to participate in the studies (Bergman, 1994).

The JCAHO also is applying outcomes data to quality management efforts in hospitals using the IMSysterm (Information Management System) (McCormick, 1990; Nadzim, Turpin, Hanold, & White, 1993). The National Committee for Quality Assurance, the organization that accredits managed care plans, has developed a tool (HEDIS—Health plan Employer Data and Information Set) for comparing managed care plans. Comparisons
involve more than 60 measures, including patient satisfaction, quality of care, and financial stability (Guadagnoli & McNeil, 1994). Researchers at the Henry Ford Health Systems’ Center for Health System Studies in Detroit have evolved 80 performance indicators to evaluate health systems (Anderson, 1991).

**Statistical Methods for Outcomes Studies**

Although outcomes researchers test for statistical significance of their findings, this determination is not considered sufficient to judge the findings as important. The focus of these researchers’ attention is on the *clinical* significance of study findings (see Chapter 12 for more information on clinical significance). In analyzing data, outcomes researchers have moved away from statistical analyses that use the mean to test for group differences. They place greater importance on analyzing change scores and use exploratory methods of examining the data to identify outliers.

**The Analysis of Change**

With the focus on outcomes studies has come a renewed interest in methods of analyzing change. Gottman and Rushe (1993) reported that the first book addressing change in research, *Problems in Measuring Change*, edited by Harris (1967), is the basis for most current approaches to analyzing change. Since then, a number of new ideas have emerged regarding the analysis of change (e.g., in studies by Collins & Horn, 1991; Rovine & Von Eye, 1991; Von Eye, 1990a, 1990b). Many researchers, however, are unfamiliar with these new ideas and continue to base their reasoning on Harris’ 1967 book. Gottman and Rushe (1993) suggest that many beliefs related to the analysis of change are based on little more than the following fallacies:

- **Fallacy 1:** In change, regression toward the mean is an unavoidable law of nature.
- **Fallacy 2:** The difference score between premeasurement and postmeasurement is unreliable.
- **Fallacy 3:** Analysis of covariance (ANCOVA, or related methods such as path analysis) is the way to analyze change.
- **Fallacy 4:** Two points (pretest and posttest) are adequate for the study of change.
- **Fallacy 5:** The correlation between change and initial level is always negative.

Outcomes researchers also are questioning the method of analysis of change. Collins and Johnston (1995) suggest that the recommended analysis method of regressing pretest scores on outcome scores and basing the analysis of change on residual change scores is overly conservative and tends to understate the extent of real change. Serious questions remain about the conceptual meaning of these residual change scores.

For some outcomes, the changes may be nonlinear or may go up and down, rather than always increasing. Thus, it is as important to uncover patterns of change as it is to test for statistically significant differences at various time points. Some changes may occur in relation to stages of recovery or improvement. These changes may occur over weeks, months, or even years. A more complete picture of the process of recovery can be obtained by examining the process in greater detail and over a broader range. With this approach, a recovery curve can
be developed, which provides a model of the recovery process and can then be tested (Collins & Johnston, 1995; Ottenbacher, Johnson, & Hojem, 1995).

**The Analysis of Improvement**

In addition to reporting the mean improvement score for all patients treated, it is important to report what percentage of patients improve. Do all patients improve slightly, or is there a divergence among patients, with some improving greatly and others not improving at all? This divergence may best be illustrated by plotting the data. For example, researchers studying a particular treatment or approach to care may develop a standard or index of various possible degrees of improvement. The index will allow better comparisons of the effectiveness of various treatments. Characteristics of patients who experience various degrees of improvement also are described, and outliers must be carefully examined. This step requires that the study design include baseline measures of patient status, such as demographic characteristics, functional status, and disease severity measures. Analysis of improvement will allow better judgments of the appropriate use of various treatments (Felson, Anderson, & Meenan, 1990).

**Variance Analysis**

Variance analysis is used to track individual and group variance from a specific critical pathway. The goal is to decrease preventable variance in process, thus helping patients and their families achieve optimal outcomes. Some of the variance is due to the presence of comorbid conditions. Keeping a patient with comorbidity on the desired pathway may require utilization of more resources early in treatment. Thus, it is important to track both variance and comorbidity. Studies examining variations from pathways may facilitate the tailoring of existing critical pathways for a specific comorbid condition.

Variance analysis also can be used to identify at-risk patients who may benefit from the services of a case manager. Variance analysis tracking is expressed through the use of graphics, and the expected pathway is plotted on the graph. The care providers plot deviations (negative variance) on the graph, allowing immediate comparison with the expected pathway. Deviations may be related to the patient, the system, or the provider (Tidwell, 1993).

**The Longitudinal Guttman Simplex Model**

The Longitudinal Guttman Simplex (LGS) Model is an extension of the Guttman scale that involves points in time, as well as items and persons. For example, an LGS model of mobility may involve the following items:

- $M_1$: moving unassisted from bed to chair
- $M_2$: moving unassisted from bed to another room
- $M_3$: moving unassisted up stairs

Table 9-7 shows hypothetical data collected with this measure on three patients at three points in time, showing a pattern of improving ability over time (Collins & Johnston, 1995).
Latent Transition Analysis

Researchers use latent transition analysis (LTA) when stages or categories of recovery have been defined and transitions between stages can be identified. To use the analysis method, the researchers place each member of the population in a single category or stage for a given point of time. However, stage membership changes over time. The analysis tests stage membership to provide a realistic picture of development. Collins and Johnston (1995) describe an example of latent transition analysis, using a hypothetical model of recovery from functional neglect after stroke, in the following excerpt:

RESEARCH EXAMPLE Latent Transition Analysis

Let’s assume that we can define a study subpopulation displaying four latent stages or types of functional neglect: sensory limitations (S), cognitive limitations (C), both (S and C) or patients may recover and adapt to the point that they are functional (F). Membership in each category is inferred from several clinical symptoms or test items, which supposedly go together but in fact may not for some patients. The items have some error and are imperfect indicators of true (latent) stage membership. Our objective is to estimate in which category a patient probably falls at any point in time and the probability of movement between stages over time, conditional on previous stage membership. Suppose we use a large number of times periodically to monitor progress, testing the same group of patients at multiple points in time. We record which items the patient passes and which the patient does not. (Collins & Johnston, 1995, p. 47)

After performing LTA by means of a computerized program designed for that purpose, the investigators presented their results in a table (Table 9-A). Data for only two points in time are shown, although the program can handle up to such five points.
Multilevel Analysis

Multilevel analysis is used in epidemiology to study how environmental factors (aggregate-level characteristics) and individual attributes and behaviors (individual-level characteristics) interact to influence individual-level health behaviors and disease risks. For example, the risk that an adolescent will start smoking is associated with the following variables: (1) attributes of the child (e.g., self-esteem, academic achievement, refusal skills), (2) attributes of the child’s

### TABLE 9-A A Hypothetical Latent Transition Model of Recovery from Neglect Following Stroke

<table>
<thead>
<tr>
<th>Latent Status</th>
<th>F</th>
<th>C</th>
<th>S</th>
<th>S and C</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total Marginal Proportions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 1 proportions</td>
<td>0.0</td>
<td>0.40</td>
<td>0.30</td>
<td>0.30</td>
</tr>
<tr>
<td>Time 2 proportions</td>
<td>0.27</td>
<td>0.25</td>
<td>0.30</td>
<td>0.18</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time 1 Latent Status</th>
<th>F*</th>
<th>C</th>
<th>S</th>
<th>S and C</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Time 1 to Time 2 Transition Proportions with Rows</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Functional (F)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Cognitive limitation (C)</td>
<td>0.46</td>
<td>0.54</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Sensory limitation (S)</td>
<td>0.30</td>
<td>0.0</td>
<td>0.70</td>
<td>0.0</td>
</tr>
<tr>
<td>S and C</td>
<td>0.0</td>
<td>0.10</td>
<td>0.30</td>
<td>0.60</td>
</tr>
</tbody>
</table>

*No patients were functional at Time 1.

The first line of the table contains the estimate of the proportion of patients in each of the four stages at Time 1. In this example, 30% of the sample had both S and C limitations, 30% had S limitations, and 40% had C limitations, and none was functional. At Time 2, the proportion in each functional limitation appears to have declined, except that for S limitations, which is unchanged, and 27% are now in the functional stage. The bottom half of the table is a matrix of transition probabilities that reveals patterns of change. Of patients who started with S, 30% improved; however, the overall percentage at S remained the same because 30% of the patients who started at S and C moved to the S category. Of patients who initially had C problems alone, 46% moved to the functional category.

A third set of quantities estimated by the full LTA model but not shown in the table are the relationships between items and stage memberships. This relationship indicates the probability that when a subject moves from one category to another, each item will also change to reflect the new stage membership. Thus, this relationship is a determination of the effectiveness of the test items or clinical symptoms as indicators of stage membership.
family (e.g., parental attitudes toward smoking, smoking behavior of parents), (3) general characteristics of the community (e.g., ease of minors’ access to cigarettes, school policies regarding smoking, city smoking ordinances, social norms of students for smoking), and (4) general social factors (e.g., geographical region, economic policies that influence the price of cigarettes). The researchers might ask, “Does smoking status covary with the level of restriction of smoking in public places after we have controlled for the individual-level variables that influence smoking risks?” (Von Korff, Koepsell, Curry, & Diehr, 1992).

**Disseminating Outcomes Research Findings**

Including plans for the dissemination of findings as a component of a program of research is a new idea within nursing if the process of dissemination is considered to be more than publishing the results in professional journals. The costs associated with dissemination are not included in funding for nursing studies beyond those of publication of the research findings. Strategies for the dissemination of research findings tend to be performed by groups other than the original researchers. The transfer of knowledge from nurse researchers to nurse clinicians has been, for the most part, ineffective.

Nursing, as a discipline, has not yet addressed the various constituencies for nursing research knowledge. A research team conducting a program of outcomes research must identify its constituencies. These should include (1) the clinicians, who will apply the knowledge to practice, (2) the public, members of which may make health care decisions on the basis of the information, (3) health care institutions, which must evaluate care in their facilities on the basis of the information, (4) health policy makers, who or which may set standards on the basis of the information, and (5) researchers, who may use the information in designing new studies. Disseminating information to these various constituencies through presentations at meetings and publications in a wide diversity of journals and magazines, as well as release of the information to the news media, requires careful planning. Mattson and Donovan (1994) suggest that dissemination involves strategies for debunking myths, addressing issues related to feasibility, communicating effectively, and identifying opinion leaders.

**KEY CONCEPTS**

- Outcomes research was developed to examine the end results of patient care.
- The scientific approaches used in outcomes studies differ in some important ways from those used in traditional research.
- The theory on which outcomes research is based was developed by Donabedian (1987). Quality is the overriding construct of the theory. The three major concepts of the theory are health, subjects of care, and providers of care.
- The goal of outcomes research is the evaluation of outcomes as defined by Donabedian, whose theory requires that identified outcomes be clearly linked with the process that caused the outcome.
- Outcomes research programs are complex and may consist of multiple studies using a variety of designs whose findings must be merged in the process of forming conclusions.
- The researcher must consider the measure’s sensitivity to change and the magnitude of change that can be detected.
• Statistical approaches used in outcomes studies include new approaches to examining measurement reliability, strategies to analyze change, and the analysis of improvement.
• Strategies must be developed in nursing to disseminate the findings from outcomes studies to the various constituencies needing the information.

TIPS FOR FURTHER STUDY

REFERENCES


status and well-being of patients with chronic conditions. *JAMA*, 262(7), 907–913.


