ASSESSMENT OF THE QUALITY OF CANCER CARE:

A REVIEW FOR

THE NATIONAL CANCER POLICY BOARD OF THE INSTITUTE OF MEDICINE

Mark A. Schuster, M.D., Ph.D.

Jennifer L. Reifel, M.D.

Kimberly McGuigan, Ph.D.

August 27, 1998

TABLE OF CONTENTS

I. INTRODUCTION

Quality assessment is becoming a major force in American health care. Consumers and purchasers of care are demanding that health plans publicly report on the quality of the care they are delivering. Local, state, and national organizations are collecting and reporting information on quality for health care organizations (most typically managed care plans), communities, and types of care.

The National Cancer Policy Board has commissioned this paper to review quality of cancer care in the United States. We will explain the methodologies and uses of quality assessment, and then report on what work has been done to assess quality for three particular types of cancer: breast cancer, childhood cancer, and prostate cancer.

We begin with a definition of *quality of care*. The Institute of Medicine has defined quality as "the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge" (Lohr, 1990).

Good quality means providing patients with appropriate services in a technically competent manner, with good communication, shared decision making, and cultural sensitivity. In practical terms, poor quality can mean *too much care* (e.g., unnecessary tests, medications, and procedures, with associated risks and side effects), *too little care* (e.g., not receiving a lifesaving surgical procedure), or the *wrong care* (e.g., medicines that should not be given together, poor surgical technique).

In recent years, the health care system has been undergoing a dramatic transformation. New organizational structures and reimbursement strategies have created incentives that could lead to a reduction in quality of care by not providing necessary care in a timely manner. Concerns about such consequences have prompted a movement to assure that efforts to control costs will not sacrifice quality.

Despite the availability of state-of-the-art health care throughout the United States, study after study has shown that many people are getting more care than they need and many are getting less than they need (Schuster et al., 1998). This is true for different types of care and in different settings.

This paper will review who can make use of information on quality of cancer care, how it can be measured, what is known about the quality of cancer care in the United States, how a health care provider can use information on quality to improve quality, and how quality is currently being monitored.

II. WHO USES INFORMATION ON QUALITY OF CARE?

Assessments of the quality of care provided by clinicians, hospitals, and health plans can be used for multiple purposes.

Information on quality can help consumers make informed choices. Most consumers know little about the technical proficiency of the clinicians or health plans they choose. Some ask friends for referrals. Some choose clinicians based on limited information in a brochure from a health maintenance organization (HMO), such as clinicians' age, gender, and medical school. Some choose providers based on convenience of location. When selecting a health plan, consumers may compare price and covered services, but they may find it difficult to learn how well plans provide care in general or for particular conditions. Quality monitoring can provide such information to help consumers decide where and from whom to obtain care.

Because a cancer diagnosis is usually unanticipated and because of restrictions on changing health insurance, most people may not specifically focus on information about the quality of cancer care when selecting a plan or source of care. Once a diagnosis of cancer has been made, many people do not have much flexibility when it comes to changing insurance. However, some people are able to change insurance (e.g., if they take a new job with a large employer for which there is no denial of coverage for preexisting conditions), and many can change medical groups within a particular health plan. There is some evidence regarding which people switch plans after a diagnosis. Although Medicare beneficiaries in managed care can change health plans or switch to fee-for-service, they do not have high rates of switching plans after a cancer diagnosis (Riley et al., 1996). However, we do not know whether these rates would be different if information on the quality of cancer care were available. In addition, it is worth keeping in mind that some people without cancer might use information on quality of cancer care (along with quality of cardiac care, diabetes care, etc.) because they want a provider who will provide excellent care for conditions for which they have a high risk (e.g., because of a particular family profile, history of smoking or occupational exposure).

Information on quality can also help public and private group purchasers of care. Companies that provide health insurance for their employees must decide which health plans to make available. Nearly

half of Americans who receive health insurance at work have only one option, so the employers are essentially choosing the plan (Mark and Burleson 1995; Etheredge 1996). Similarly, state and federal governments must decide which health professionals can provide care to beneficiaries of government-funded health insurance (e.g., Medicaid, Medicare). As concerns have increased about rising health care costs, many large purchasers have considered price of care as the primary factor in determining which health plans to offer. Yet, although it is tempting to believe that more efficient plans trim costs in health care while keeping essential and necessary services, studies reveal that cost containment results in the elimination of both necessary and unnecessary care (Bernstein et al., 1993; Hilborne et al., 1993; Leape et al., 1993; Lohr et al., 1986). Quality assessment and monitoring provide the tools purchasers need to balance cost and quality. Some organizations representing purchasers of care (e.g., the Pacific Business Group on Health) incorporate quality improvement goals into their negotiations over price. If the plans do not meet the goals, they forfeit a portion of their premiums.

Information on quality is also useful for physicians and patients when making specific treatment and referral decisions. For example, success rates for a particular intervention may be highest in research studies, which often include only the best clinicians and facilities and enroll patients who have no other diseases and who are likely to follow treatment instructions carefully (Brook, 1993). Because care provided in the "real world" may not have the same results as care provided under research conditions, information on local clinical results will help primary care providers and patients determine which available physicians or hospitals have the best success rates and, given those rates, whether the potential benefits of a procedure found effective in a research setting outweigh the risks. For example, in the 1980s, single institution studies found higher response rates with new chemotherapy regimens (m-BACOD, ProMACE-CytaBOM, MACOP-B) compared with the standard treatment for Non-Hodgkin's Lymphoma (CHOP). When multi-institutional studies were conducted, the success rates were not replicated, presumably because the original studies used highly select samples of patients (Gordon et al., 1992; Fisher et al., 1993).

Clinicians and health plans can use information concerning their own quality of care to improve the care they provide. Monitoring quality provides the opportunity not only to address inferior quality, but also to identify and learn from examples of superior quality. Health plans may find that they need to help individual clinicians improve their care, or they may identify organizational problems that hinder provision of high quality care. Clinicians and plans may welcome the opportunity to improve, and those that might not otherwise address quality problems may be motivated by public reporting to address them.

The health system in the United States is rapidly changing because of private market forces and public policies. However, we know relatively little about the impact of these changes on the health of the population. Sometimes we cannot even predict whether a change will increase or decrease quality because the change creates multiple, competing incentives for the various participants in the system. With quality measurement systems, we can evaluate the impact of health system changes. For example, Medicaid beneficiaries in many states are being enrolled in managed care organizations. To determine what impact this shift has on beneficiaries' health, measurement tools could be used to compare the quality of care patients with cancer (including stage at diagnosis) receive in managed care settings and the quality of care they receive in traditional fee-for-service settings.

Cost-cutting efforts create incentives for health plans to reduce the amount of care provided, which may actually improve the health of the population when useless or harmful care is eliminated. However, when necessary care is cut, health may decline. Quality of care criteria can help guide decisions about which type of care should be maintained and which type might be safely eliminated. Better decisions should result in the delivery of a more effective package of services. There has been some concern, though, that cost-cutting efforts can be introduced under the guise of quality improvement, without real attention to improving quality. Such efforts should not be allowed to undermine sincere efforts at making quality better.

III. HOW DO WE MEASURE QUALITY?

Quality assessment is the measurement of quality by expert judgment (called implicit review) or by systematic reference to objective standards (called explicit review). Quality may be evaluated at any level of the health care system: for physicians and other health care professionals; for hospitals, clinics, rehabilitation centers, and other institutions; for health plans; and for communities.

Different approaches to assessing quality have different strengths and weaknesses, and some approaches work better in one setting than another. An example of implicit review is when a clinician

reviews the medical record of a patient and expresses a judgment on whether the care was good or bad. The clinician may base her opinion on years of experience and understanding of the clinical situation for which care was provided. However, she might not give the same rating on a different day, and her colleagues might give a different rating.

Explicit review provides a more systematic approach and can be based on one or more of three dimensions: structure, process, and outcomes (Donabedian, 1980). "Structural quality" refers to health system characteristics, "process quality" refers to what the clinician does, and "outcomes" refers to patients' health. While producing good outcomes is the ultimate goal of the health care system, using outcome measures to assess quality is not usually the most effective approach for a variety of technical reasons (which we discuss below). Therefore, we often use process measures instead. In this section, we review each type of measure and how it is used.

A. Structural Quality

Structural quality refers to characteristics of the health system that affect its ability to meet the health care needs of individual patients or communities. These characteristics include clinician characteristics (e.g., percentage of physicians who have board certification, average years of experience, distribution of specialties), organizational characteristics (e.g., staffing patterns, reimbursement method), patient characteristics (e.g., insurance type, illness profile), and community characteristics (e.g., per capita hospital beds, transportation system, environmental risks). Structural measures specifically related to cancer quality could include presence of a multi-disciplinary cancer center, a bone marrow transplant unit, or psychological support services.

Structural characteristics are often necessary to provide good care, but they are usually insufficient to ensure excellent quality. The best structural measures are those that can be shown to have a positive influence on the provision of care (process quality) and on the patients' health (outcomes), though this relationship has not been found for most measures (Brook et al., 1990). One structural measure that has been found to relate to outcomes for some conditions or procedures is volume, which refers to the number of times each year that a clinician (or hospital) performs a particular procedure or takes care of patients with a particular disease. Several studies have found that the higher the volume, the better one's results (Grumbach et al., 1995; Hannan et al., 1997; Kitahata et al., 1996; Luft et al., 1990). There is also evidence that higher volume is related to better outcomes for cancer care. One study from Scotland showed that centralized treatment improves outcomes for non-seminomatous germ-cell tumors, even when controlling for participation in a protocol (although protocol treatment explained a large part of the variation) (Harding et al., 1993). Some pediatric cancer studies have also looked at this issue and will be discussed below in the section on childhood cancer.

Structural quality measures have long been the key component in accreditation procedures. Various independent organizations accredit hospitals or health plans based on a set of criteria that generally focus on structural measures such as appropriate capacity for the covered patient population. In recent years, accreditation organizations have also been incorporating process and outcome measures into their accreditation procedures.

B. Process Quality

Process quality refers to what clinicians do for patients and how well they do it, both technically and interpersonally.

1. Types of process quality

a. Technical process refers to whether the right choices are made in diagnosing and treating the patient and whether care is provided in an effective and skillful manner. The former is generally measured by assessing (i) appropriateness or (ii) adherence to professional standards, which we describe below. The latter is more difficult to assess directly (e.g., it may require direct observation of the clinical interaction) and may be inferred from outcomes (e.g., complications following surgery).

i. Appropriateness. An intervention or service (e.g., a laboratory test, procedure, medication) is considered appropriate if the expected health benefits (e.g., increased life expectancy, pain relief, decreased anxiety, improved functional capacity) exceed the expected health risks (e.g., mortality, morbidity, anxiety anticipating the intervention, pain caused by the intervention, inaccurate diagnoses) by a wide enough margin to make the intervention or service worth doing (Brook et al., 1986). A subset of appropriate care is *necessary* or crucial care. Care is considered necessary if there is a reasonable chance of a nontrivial

benefit to the patient and if it would be improper not to provide the care. In other words, this is care that is important enough that it might be considered ethically unacceptable not to offer it (Kahan et al., 1994; Laouri et al., 1997). Appropriateness criteria and necessity criteria can be used, respectively, to measure *overuse* of care, which occurs when expected risks exceed expected benefits (which is a problem because of treatment complications and wasted resources), and *underuse* of care, which occurs when people are not getting care that is expected to improve their health.

Studies of appropriateness of care show that both underuse and overuse are common and can exist in the same community or institution. People with a particular disease can get too much of one type of care and too little of another type. This may be true of cancer, although it has not been studied.

ii. Adherence to professional standards of care. Another way to measure process quality is to determine whether care meets professional standards. This assessment can be done by creating a list of quality indicators that describe a process of care that should occur for a particular type of patient or clinical circumstance. Quality indicators are based on standards of care, which are found in the research literature and in statements of professional medical organizations or which are determined by an expert panel. The performance of physicians and health plans is assessed by calculating rates of adherence to the indicators for a sample of patients. Current performance can be compared to a physician's or plan's own prior performance, to the performance of other physicians and plans, or to benchmarks of performance. Indicators can cover a specific condition (e.g., patients diagnosed with colon cancer who do not have metastatic disease should be offered a wide resection with anastamosis within six weeks of diagnosis) or they can be generic, covering general aspects of care regardless of condition (e.g., patients prescribed a medication should be asked about medication allergies).

b. Interpersonal quality refers to whether the clinician provided care in a humane manner consistent with the patient's preferences. It includes such topics as whether the clinician supplied sufficient information for the patient to make informed choices and involved the patient in decisionmaking. It is generally assessed using patient survey data.

2. Attributes of Good Process Measurement

Good process measures are supported by research studies or professional consensus. They are also flexible with respect to patient preferences. Some patients may not want what most people would consider proper care. Indicators can be constructed so that they are scored favorably if the care was offered but declined. However, there needs to be some recognition that a perfect score on indicators is not necessarily a feasible or even a desirable goal. For example, although chemotherapy is highly recommended after surgical resection of colon cancer involving the lymph nodes, some patients might decline treatment because they do not wish to experience its associated toxicities. Therefore, 100% adherence may not be a reasonable target for an indicator specifying adjuvant chemotherapy for these patients. Furthermore, such a target might also create incentives to ignore patient preferences in making treatment decisions. An alternative approach would be for an indicator to specify that chemotherapy was offered or recommended.

The best process measures are those for which there is evidence from research that better process leads to better outcomes. For example, adjuvant chemotherapy has been shown in several randomized controlled trials to improve survival after surgery for Duke's C colon cancer (NIH, 1990); performing routine mammography identifies breast cancer at an earlier stage when it is more curable (Kerlikowske et al., 1995); perioperative chemotherapy and radiation therapy have been shown to increase survival for patients with rectal cancer (Krook et al., 1991; Moertel, 1994). Unfortunately, research has not covered all aspects of standard medical practice related to cancer (or other types of disease), and so in those cases we rely on expert consensus to decide which processes are important measures of quality. If there is not strong consensus supporting the value or superiority of a clinical practice, it is usually not used as a quality measure.

Several studies outside of oncology have tied process measures to outcomes. In a study of five hospitals in Los Angeles County, mortality rates were examined for patients who had coronary angiography and for whom a revascularization procedure was deemed necessary by explicit criteria. Those who received necessary revascularization within one year had a mortality of 9%, compared to 16% for those who did not. Those receiving necessary revascularization also had less chest pain at follow-up (Kravitz et al., 1995).

Other research also demonstrates the link between process and outcome. In a study of Medicare enrollees hospitalized with congestive heart failure, heart attack, pneumonia, and stroke in 1981-82 and 1985-86, better process quality of care was significantly associated with lower mortality rates 30 days after hospitalization. Patients who went to hospitals in the lowest 25th percentile on a set of process of care measures had a 39% increased likelihood of dying within 30 days after hospital admission compared to patients who went to hospitals in the highest 25th percentile, after adjusting for patient sickness at admission (Kahn et al., 1990).

C. Outcomes

Measurement of outcomes is probably the most intuitively appealing approach to quality monitoring. "Outcomes" refers to the results of a health care delivery process. The three main types of outcomes are clinical status, functional status, and consumer satisfaction.

1. Types of outcome measure

a. Biological status.

Biological status consists of the biological outcomes of disease, such as how organ systems are functioning. Physicians have traditionally used biological status to determine treatment success or failure. Cancer research, for example, has long used the outcome of five year overall survival or five-year progression-free survival. Other measures might include post-operative wound infections or catheter infections. Proxy measures (sometimes called surrogate end points or intermediate outcomes) are also used. They do not measure the outcome of concern directly, but they provide evidence or likelihood of a good outcome. For example, response rate (decrease in tumor size) is used to assess the impact of therapy, but the goal of therapy may be prolonged life. When used as a measure of quality or as an indicator of impact of therapy, it is important for there to be evidence that the proxy measures are really serving as a proxy. In other words, the effect of the intervention on the proxy should be concordant with the effect on the cancer itself (Schatzkin et al., 1996).

b. Functional status.

Measures of functional status assess how disease affects an individual's ability to participate in physical, mental, and social activities. It also covers the ability to meet the regular responsibilities of one's roles in society (e.g., parent, bank teller, volunteer). Health-related quality of life is similar to functional status but also includes the person's sense of well-being as well as factors external to the individual such as social support.

Functional status assessment is based on the premise that many aspects of health are important to patients and will influence them in making treatment decisions. Such assessment could help someone choose between a treatment that would give many more years of life with major incapacitation and a treatment that would give fewer years of life with full function. For example, treatment success/failure for prostate cancer has historically been assessed by the clinical outcome of whether the patient died from prostate cancer. However, functional status measures would incorporate other treatment outcomes, such as the patient's urinary, sexual, and bowel function (Litwin et al., 1995). Functional status assessment often includes the degree to which disease limits one's ability to participate fully in activities of daily living. Depending on the type of cancer and phase of illness, such activities could include going to work or caring for children. In a patient with more advanced disease, however, an assessment of whether they are able to go to the market for groceries or bathe or dress themselves may be more relevant.

Performance status is a measure of functional status often used in oncology clinical trials. The Karnofsky Performance Status is a rating of patients' functional status that has been used in clinical trials since 1949 (Karnofsky and Burchenal, 1949; Greico and Long, 1984). The rating is performed by a physician or nurse. It has been found to be a strong predictor of survival in some patient populations, most notably patients with lung cancer. However, it covers only one aspect of quality of life—physical performance—and, although significantly correlated with quality of life, it accounts for less than 50% of the variability in patients' own ratings of their quality of life. While clinician-rated measures have value, the field is moving more towards use of patients' assessments of functional status and quality of life (Reifel et al., 1998), which are preferable model for quality assessment. Examples of patient-based measures include

the Cancer Rehabilitation Evaluation System (CARES) (Coscarelli Schag and Heinrich, 1990; Ganz et al., 1992), the Functional Living Index-Cancer (FLIC) (Schipper et al., 1984), and the Breast Cancer Chemotherapy Questionnaire (BCQ) (Levine et al., 1988).

c. Consumer satisfaction.

Consumer satisfaction refers to how patients feel about the care they receive and is generally measured by surveys of patients. There is a relationship between satisfaction and adherence to treatment regimens. Patients who are satisfied are more likely to take their antibiotics properly (Bartlett et al., 1984), follow treatment recommendations (Hsieh and Kagle, 1991), and return for follow-up visits (Deyo and Inui, 1980). Thus, the physician has an incentive to please his/her patients as part of the treatment—so they will be more likely to follow his/her advice. Furthermore, dissatisfaction with care can lead patients to switch clinicians and health care institutions (Reichheld, 1996; Rubin et al., 1993; Young, 1985).

One limitation of satisfaction ratings is that consumers cannot always tell if the care was appropriate or technically good (Aharony and Strasser 1993); research has not shown a consistent relationship between consumer satisfaction and technical quality of care (Cleary and McNeil, 1988; Davies and Ware, 1988; Hayward et al., 1993). A kind and caring physician may provide care that is technically poor (Aharony and Strasser, 1993). Also, consumer satisfaction may vary with expectations. For example, patients who have a history of poor access to health care may be so appreciative when they actually see a physician that they may report high satisfaction regardless of how well care was delivered. Therefore, it is best to use satisfaction ratings as one component of quality assessment rather than as the sole form of measurement.

2. Treatment Effects vs. Quality of Care

Biological and functional status can be measured for more than one purpose. We have described them here in the context of quality of care, in which the outcomes are compared between two institutions as a sign of whether one institution is delivering better care (with the presumption that better care leads to better outcomes). However, these measures are also used clinically to track a patient's progress and in clinical trials to measure the efficacy or effectiveness of a new drug or intervention. The same measures can sometimes be used for both purposes, but certain measures are better suited for one purpose or the other. Five-year survival rates, for example, are a standard measure used in studies of new cancer treatments. However, when measuring quality of care for purposes of accountability or quality improvement, we generally need to use outcomes with a shorter time horizon than five years. If we compared two institutions using five-year survival rates for colon cancer, we might find that one institution had higher survival rates. However, in the interim, there might have been a change in staff or a revamping of procedures that improved or weakened the quality of care at the hospitals, thereby making the comparison of historical but not practical value.

3. Attributes of Good Outcomes Measurement

Outcomes measurement has become increasingly popular in the past few years, perhaps because outcomes are the most direct measure of the health of a population. For example, outcomes can be used to assess the quality of care a health system provides its cancer patients: outcomes can measure whether the survival and quality of life of women diagnosed with breast cancer and whether they are satisfied with their care.

The best outcome measures have certain key features or are used in a particular manner. First, they should be risk-adjusted (or case-mix adjusted), in other words, adjusted for factors that influence outcomes but are beyond the health care system's control (e.g., age, socioeconomic status, co-morbidities). Without such adjustment, it is impossible to determine how much of the improvement or worsening of outcomes is due to the care delivered (or not delivered) by the health care system. A radiation oncologist who receives referrals of patients with multiple medical problems is likely to have worse outcomes than one who takes only patients with mild disease, even though the former may be a better radiation oncologist. To make comparisons in such a system, we need to adjust for how ill the patients are. Risk-adjustment is complex, and the selection of factors to use in risk-adjustment must be done carefully to allow for accurate interpretation of the outcomes (Iezzoni, 1996).

Outcomes are most useful when we know the specific processes of care that relate to them. Then, if the outcomes are not as good as they should be, we know what aspects of care to address to try to

improve them. In other words, if we do not know how an outcome relates to processes of care, we will not know what to do to improve the outcome when we find that it is poor at a particular hospital.

It also helps to measure outcomes from different perspectives. For example, palliative chemotherapy for metastatic cancer may decrease a patient's tumor burden and potentially prolong her life, but it might also cause severe fatigue and weight loss, so the patient's clinical status might improve while functional status declines.

It is also important to use outcomes that can be reasonably related to the health care system one is assessing and to the particular part of the system one is assessing. Sometimes a single outcome may be influenced by many different clinicians over many years. For example, outcomes for lung cancer may reflect the quality of care provided over many years, including the quality of smoking prevention counseling for adolescents and smoking cessation counseling for adolescents and adults. Outcomes for breast cancer may in part depend on the quality of screening and early detection. Given how frequently most patients change clinicians or health plans, it could be difficult to relate the quality of any one clinician or plan to some outcomes. Similarly, if one is trying to use outcomes to assess the quality of surgeons treating a sarcoma at various hospitals, it is important to distinguish whether the outcomes are related to the skill of the surgeon, the competence of the surgical team, or organizational characteristics of the hospital. One might also want to consider the skill of the medical oncologist prescribing neoadjuvant chemotherapy. For breast cancer, treatment may depend upon an oncologist, a surgeon, and a radiation oncologist. It can be difficult to distribute responsibility among them.

In addition, outcomes should be measured on samples that are large enough to detect differences in quality. Adverse outcomes are often uncommon events, so large samples are needed to detect clinically meaningful differences between hospitals. To detect a two percentage point difference in rate of catheter infections between two hospitals (e.g., five percent for one and seven percent for the other), each hospital would need to have at least 1900 patients with a catheter.

In summary, there are many challenges inherent in using outcomes to measure quality of care. If we do not address them, we may find that we measure different outcomes among the patients of three physicians but that we cannot determine whether those different outcomes are attributable to the physicians. Process measures have their own challenges (e.g., one needs to make sure that the processes really are important for improving outcomes), but such measures can be quite effective in showing whether providers are doing what they should so that their patients have the best chance of achieving good outcomes. We have more experience using process measures than outcome measures to assess quality, and many quality assessment systems depend primarily or exclusively on process measures. However, interest is growing in improving outcomes measurement, so that outcomes can be used along with process measures to provide more useful assessments of health care quality.

D. How is Quality of Care Information Collected?

Data for quality assessment can come from several sources. First, administrative records are widely available, and while they are limited in clinical detail, they can be used to show intensity or patterns of utilization. For example, they can be used to determine whether a patient with large cell NHL received at least six months of chemotherapy and whether the patient had a white blood cell count performed before receiving chemotherapy. Second, medical records can provide greater clinical detail, such as what was recorded from the medical history, what the results of laboratory tests were, and what the treatment plan was. For example, the medical record can show whether the patient with NHL was neutropenic, whether proper components of the physical examination were performed, and whether chemotherapy was held. Third, patient surveys can provide additional useful information. Patients can report on what happened during a clinical encounter and thereby provide information relevant to the processes of care. They can also rate their satisfaction with care and provide information on outcomes such as functional status. It is generally more expensive and time-consuming to collect information from medical records and from patients than from administrative data.

Cancer registries are also a potential source of information. They collect information on type of cancer, histology, stage at diagnosis, patient age, and initial course of treatment (whether the patient received surgery, chemotherapy, and radiation therapy that would normally be prescribed as part of the initial treatment plan). Registries exist at the state, national, and international level.

There are two main national registries: the Surveillance, Epidemiology, and End Results (SEER) Program of the National Cancer Institute (NCI) and the National Cancer Data Base (NCDB). The SEER program was established as a result of the National Cancer Act of 1971 in order to assemble, analyze, and

distribute information on the prevention, diagnosis, and treatment of cancer. The program routinely collects information from designated population-based cancer registries from different parts of the country. The different areas have been chosen for their capacity to maintain a cancer reporting system as well as for their ability to report epidemiologically significant population subgroups. Currently, 13.9% of the U.S. population is represented by the geographic areas that make up the SEER Program's data base. Some goals of the SEER program include compiling (with the help of the National Center for Health Statistics) estimates of cancer incidence and mortality in the U.S.; discovering trends and unusual changes in specific cancers based on their geographic, demographic, and social characteristics; as well as providing information about trends in therapy, changes in the extent of disease (stage at diagnosis), and changes in patient survival; and promoting studies that identify those factors that can be controlled through intervention strategies.

The National Cancer Data Base (NCDB) is a joint project of the Commission on Cancer (COC) of the American College of Surgeons (ACoS) and the American Cancer Society (ACS) to facilitate community, hospital, state, and national assessment of care of patients with cancer (Menck et al., 1997). It began in 1988, and in 1994, 1227 hospitals reported data on 689,714 cases to the NCDB, which is estimated to be 57% of new cancer cases that year. NCDB collects information on patient characteristics, tumor characteristics, first course of treatment, and follow-up. Participating hospitals submit all cases seen at their hospital for a particular data year. The system appears to have a bias towards hospitals with a computerized cancer registry, and it does not provide comprehensive outpatient data. After 1996, all approval hospitals (i.e., hospitals with ACoS accreditation, about 1450 hospitals) were required to participate; it is estimated that in 2000, 1750 hospitals representing 80% of U.S. cancer cases will be participating. The NCDB provides comparisons of cancer management patterns and outcomes to national norms at the hospital, community, and state levels. The NCDB can also be used to track how well results of major clinical trials are incorporated into clinical practice. There have been questions about the quality control of the data collected by the individual hospital registries, and NCDB is working to improve the quality.

There have been questions raised about the accuracy of data reported to registries, but we found no published studies on the subject. To address these questions, the UCLA Division of Cancer Prevention and Control Research is conducting a study in colloboration with PacifiCare, a managed care plan, to determine the reliability of the treatment information collected by the California Cancer Registry.

Registries represent an exceptionally valuable opportunity to conduct quality assessment on a broad level. They could go further in collecting data on explicit process measures as well as on characteristics needed to risk-adjust the outcomes. They would also be more useful for quality assessment if they were able to reduce the time lag between provision of care and availability of data. This information could be used both to provide detailed information on quality of care as well as to tie the processes of care to the outcomes. These data would also be useful in quality improvement.

E. What Does Quality of Care Information Look Like?

Quality of care information is often made available in the form of report cards. Report cards use a standardized format for making results of quality measures available to consumers (e.g., when choosing among health plans), purchasers of care, physicians, health plans, and health care institutions. They generally show comparative scores on quality measures for different clinicians or entities on single quality measures or groups of measures. Report cards are most useful when standardization of information allows reasonable and fair comparisons among clinicians or plans and when they present results in a way that does not mislead users into placing undue importance on differences that are not statistically significant or clinically relevant.

IV. HOW CAN AN ORGANIZATION USE INFORMATION ON QUALITY TO IMPROVE THE QUALITY OF CARE IT PROVIDES?

Quality assessment conducted by an organization (e.g., a private practice, hospital, health plan) for internal use is called internal quality assessment. Ideally, information on quality shows where to focus efforts to improve quality. Even when the information is collected so that another organization can conduct external quality assessments of multiple plans or providers (see next section), it can generally be adapted for use in internal quality improvement.

Several techniques are available for using internal quality assessments to improve quality within an organization. Traditional quality assurance programs focus on improving or removing clinicians who

stand out as performing below group norms. By contrast, continuous quality improvement (CQI) programs (also known as total quality management or TQM programs) focus on improving the quality of care delivered by all clinicians, with the goal of raising the average level of quality in an organization. CQI assumes that most examples of poor quality are due to correctable systematic problems rather than to individual incompetence or irresponsibility. It often incorporates routine collection and monitoring of information to assess quality so that the organization can identify and respond to problems in a timely fashion. CQI generally uses interdisciplinary teams both to identify inefficiencies that increase errors and to institute checks that make errors easier to prevent and catch (Berwick, 1989; Kritchevsky and Simmons, 1991). For example, a group of all cardiothoracic surgeons practicing in three states used CQI and other techniques to improve their practices and found a 24% reduction in their combined mortality rates (O'Connor et al., 1996).

A hospital that finds many medication errors might review the many steps involved in the process by which physicians, nurses, pharmacists, and orderlies provide medications to patients. Solutions might include having nurses and pharmacists double-check doses, using standardized doses for patient weight ranges to reduce calculation errors, and having a standard location for documentation of all drug allergies (Leape et al., 1995). Prevention of medication errors is particularly important for cancer chemotherapy, and efforts have been initiated to improve safeguards (Cohen et al., 1996).

One approach a hospital might take if it finds wide variation in how a clinical condition is being treated is to provide practice guidelines, which are a set of systematically-developed recommendations about some or all aspects of decisionmaking for a particular condition or clinical situation (Field et al., 1990). The development of guidelines involves a review of the relevant research and clinical literature, and the best guidelines incorporate citations to the literature. Although guidelines can be presented as a simple list, they are often structured as a decision algorithm with branch points and if-then statements, thereby allowing clinicians to understand the rationale behind each step involved in diagnosing and treating the condition. Guidelines are not intended to dictate a rigid approach to care; rather, they give options that a clinician should be aware of, even if he or she chooses a different strategy for a particular patient. They are partly an outgrowth of the boom in scientific information that makes it difficult for individual physicians to keep up with medical advances. Practice guidelines have been developed by government agencies (e.g., the Agency for Health Care Policy and Research), specialty organizations (e.g., American Society of Clinical Oncology), and cancer centers (e.g., the National Comprehensive Cancer Network).

V. WHAT SYSTEMS ARE USED FOR EXTERNAL QUALITY MONITORING IN THE UNITED STATES?

External quality assessment makes quality information routinely and systematically available to consumers and large purchasers of care to inform their choices of physicians, hospitals, and health plans. Without such quality comparisons, variations in the costs of plans may not be meaningful.

In this section, we discuss several examples of different types of organizations and systems that are measuring quality of care in the U.S. We also describe the ways in which these systems incorporate information on quality of cancer care.

A. Health Plan Employer Data and Information Set

The non-profit National Committee for Quality Assurance (NCQA) produces a widely-used report card monitoring system called the Health Plan Employer Data and Information Set (HEDIS). HEDIS is a performance measurement tool designed to assist purchasers and consumers in evaluating managed care plans and holding plans accountable for the quality of their services. Because HEDIS has standard measures and uniform data reporting requirements, comparisons can be made across various health plans and their organizational structures (e.g., staff model health maintenance organizations, point of service plans).

The most recent iteration, HEDIS 3.0, assesses plans in eight domains: effectiveness of care, accessibility and availability of care, satisfaction with care, cost of care, stability of the health plan, informed health care choices, use of services, and plan descriptive information. In addition to the standard set of measures, there is a set of measures that are not used in current scoring but are being tested for use in future iterations of HEDIS. Several indicators are related to cancer (See Table 1).

HEDIS is a voluntary system, although managed care plans are finding it increasingly necessary to participate to compete for patients. Most large employers request HEDIS data from the managed care plans with which they contract. In 1996, more than 330 plans—over half the plans in the U.S. representing more

than three-quarters of all commercial managed care enrollees—were reporting HEDIS measures on their commercial enrollees.

NCQA produces Quality Compass, a CD-ROM-based system that makes it possible for consumers to obtain comparative HEDIS ratings for HMOs in communities throughout the U.S. In 1997, the system covered health plans representing three quarters of all HMO-enrollees in the U.S. A subset of Quality Compass measures appears on the World Wide Web.

B. Foundation for Accountability

The Foundation for Accountability's (FACCT) principal mission is to ensure that information on quality is effectively communicated to and used by consumers. FACCT has developed a quality assessment framework that uses categories that are designed to be descriptive of issues consumers care about: The Basics, Staying Healthy, Getting Better, Living with Illness, and Changing Needs. FACCT has been developing new measures and selecting existing measures for use by others in assessing quality.

One of the areas for which FACCT has been compiling a set of measures is breast cancer (See Table 1). The indicators are being evaluated in several sites around the country.

C. Consumer Assessment of Health Plans Study (CAHPS)

The Consumer Assessment of Health Plans Study (CAHPS), funded by AHCPR, has developed surveys that ask consumers about their experiences and satisfaction with receiving health care. A sample of people enrolled in each health plan fills out the surveys. CAHPS provides a reporting system designed to provide information about health plans (including consumers' experiences with receiving health care) for the benefit of people in the process of selecting a plan. CAHPS does not focus on specific diseases but rather asks about general aspects of care.

D. Pacific Business Group on Health (PBGH)

PBGH is a nonprofit coalition of large health care purchasers in California and Arizona created in 1989 to improve the quality of health care and address rising costs. As of October 1996, it represented 33 private and public sector organizations, almost all of which have at least 2000 benefits-eligible employees. PBGH estimates that it represents about 2.5 million insured individuals. PBGH members have established a Quality Improvement Fund that supports projects to measure and improve quality provided by health plans, hospitals, and physician groups.

One of its projects is the California Collaborative Healthcare Reporting Initiative (CCHRI), which represents the collective interests of health plans, purchasers, and provider organizations. CCHRI collects, analyzes, and reports performance data; promotes the use of accurate and comparable quality measures; and designs and engages in collaborative quality improvement projects. It provides purchasers with this information to gauge health plan performance, and it makes this information available to the public in a report card format.

In 1994, employers that belong to PBGH formed a Negotiating Alliance to negotiate collectively with California managed care organizations (MCOs) to establish a standardized benefits package emphasizing both price and quality. PBGH negotiates an individual set of performance targets with each participating MCO. If the MCO does not meet the targets, then 2% of the premium paid by PBGH employers to the MCO is turned over to PBGH. Therefore, the MCOs have a financial incentive to meet their targets.

PBGH is also developing several disease-specific quality assessment programs, including one for breast cancer.

E. Joint Commission on Accreditation of Healthcare Organizations (JCAHO)

The Joint Commission on Accreditation of Healthcare Organizations (JCAHO) is one of the major accrediting organizations in the U.S. It is a non-profit organization that evaluates and accredits more than 16,000 health care organizations in the U.S.

For accreditation, JCAHO conducts an on-site quality assessment every three years. It covers such topics as patient rights, patient care, patient education, continuity of care, ongoing efforts to improve quality, safety plans, information management, and infection control. Although JCAHO (and other accrediting organizations) traditionally focused on structural measures of quality, such as whether a hospital has appropriate capacity for the covered patient population, it now incorporates process and outcome measures into its accreditation criteria.

JCAHO has instituted the ORYX system, whereby organizations seeking JCAHO accreditation need to select a performance measurement system and two specific indicators on which they will be reporting on their care. Hospitals and long term care facilities are being instructed to begin reporting on these indicators during early 1999. One option is to use JCAHO's Indicator Measurement System, which has specifications for 42 quality of care indicators and is developing others. Five of these indicators cover cancer care (See Table 1).

Self-selection of indicators will make comparisons across institutions challenging, but it should allow for comparisons to prior years, benchmarks, and goals.

Application for JCAHO accreditation is voluntary. About 80% of U.S. hospitals participate; they represent about 96% of all inpatient admissions. JCAHO provides different levels of accreditation. In 1995, 12% of hospitals received accreditation with commendation, 10% received accreditation, 74% received accreditation with recommendations for improvement (which had to be met within a specified period of time), 3% received provisional accreditation (only hospitals newly seeking accreditation are eligible for this category), less than 1% received conditional accreditation (indicating serious problems that needed to be addressed), and less than 1% received no accreditation. A new category of preliminary nonaccreditation was added in 1996 to cover the initial period of accreditation denial.

JCAHO accredits not only hospitals and health plans, but also home health care organizations, clinical laboratories, and other health care facilities.

VI. BREAST CANCER CASE STUDY

Multiple characteristics of breast cancer make it the model oncologic condition for quality of care evaluation. Breast cancer is the most common neoplasm affecting American women, and it is estimated that one in eight women will develop breast cancer in her lifetime (Ries et al., 1998). In 1998, 178,800 new breast cancer diagnoses are expected (Landis et al., 1998). While lung cancer has surpassed breast cancer as the leading cause of cancer deaths in women, breast cancer still accounts for nearly 43,500 deaths annually (Landis et al., 1998). Breast cancer remains a common illness with significant morbidity and mortality.

A. Breast Cancer: Extensive Literature Supports Process-Outcomes Links

As discussed in the introduction, the best process measures are those for which there is evidence from research that better process leads to better outcomes. Among oncologic conditions, breast cancer has one of the most extensive scientific literatures to support a strong association between processes of care and outcomes. Unlike many malignancies, effective interventions exist for breast cancer that decrease mortality and improve quality of life. In addition, the evidence from the literature suggests that all phases of the continuum of care have an important effect on breast cancer outcomes, including early detection, diagnostic evaluation, and treatment. This extensive clinical literature with many well-designed randomized controlled trials provides a firm grounding for the development of process measures in breast cancer. However, even in breast cancer, not every aspect of the continuum of care has been studied to determine its effect on outcomes. Thus, even in this heavily studied disease, some of the presumed associations between process and outcomes reflect consensus within the medical community and expert opinion and are not the results of randomized controlled trials. To further illustrate these points, we will briefly review the literature on early detection, diagnostic evaluation, and treatment of breast cancer, with emphasis on where process-outcomes links can be based on evidence and where they must be based on expert opinion and consensus.

The first phase of the breast cancer care continuum, screening, has ample evidence in the literature for process-outcomes links. Although there has been controversy over the age at which one should begin screening for breast cancer, data from multiple randomized trials and several meta-analyses provide evidence that screening with mammography results in diagnosis at an earlier stage and better breast cancer outcomes. Early detection of breast cancer through screening mammography has been shown to reduce mortality by 20% to 39% for women ages 50 years and older (Shapiro et al., 1988; Nyström et al., 1993; Roberts1990). And, while the results for women 40 to 50 have not reached statistical significance, studies have shown a trend toward reduced mortality ranging from 13% to 23% (Kerlikowske et al., 1995; Tabar et al., 1995). Many professional and public health organizations, including both the American Cancer Society and the National Cancer Institute, currently recommend screening mammography beginning at age 40

(Mettlin et al., 1994; Eastman, 1997). The evidence for efficacy of screening mammography, as well as the consensus among the medical community, provide the support for a strong process-outcomes link for routine mammography and breast cancer outcomes. It is then perhaps not surprising, as we will discuss in a later section, that one of the first performance measures included in HEDIS and other measurement sets is rates of screening mammography.

The next phase along the continuum of care for breast cancer, the diagnostic evaluation of a breast abnormality discovered either through physical exam or mammography, is important for the timely diagnosis of breast cancer and therefore for patient outcomes. However, research on the evaluation of a breast abnormality primarily addresses the issue of diagnostic accuracy and does not provide direct evidence of a process-outcomes link. Nonetheless, widespread consensus dictates that a woman presenting with a breast mass should have it evaluated to be sure that it is not a malignancy (Osuch, 1996). The initial approach to a breast mass varies with the patient's age and the characteristics of the mass on physical exam. However, any persistent breast mass that is not determined to be a simple cyst either by aspiration of clear fluid or by ultrasound characteristics should have a pathologic diagnosis, either through fine needle aspiration or excisional biopsy. Biopsy is the gold standard. Fine needle aspiration has a false negative rate ranging from 0.4% to 35%, but when used in conjunction with clinical exam and simultaneous mammography in a procedure referred to as triple diagnosis, the positive predictive value of a negative finding for all three is 99% (Kaufman et al., 1994).

Similarly, the impact of the follow-up of an abnormal mammogram on breast cancer outcomes has not been studied directly. However, in order for screening mammography to be effective, abnormalities identified at screening must be evaluated appropriately. While certain mammographic images (e.g. a spiculated mass) are characteristic of cancer, no criteria allow the radiologist to absolutely differentiate benign from malignant lesions (ACOG, 1994; Osuch, 1996; Talamonti and Morrow, 1996). Given this significant overlap in the appearance of benign and malignant lesions on mammography, other tests are necessary to rule-out a malignancy. Ultrasound may be useful to differentiate cysts from solid masses; however, a biopsy to obtain a pathologic diagnosis is often the only way to determine if a lesion is benign or malignant.

These studies do not specifically link effective diagnostic work-up for a breast mass or an abnormal mammogram to eventual breast cancer outcomes. However, expert opinion and guidelines from a number of professional organizations endorse the processes described for the evaluation of a breast abnormality (Talamonti and Morrow 1996). In addition, timely and appropriate evaluation of breast abnormalities is also considered to be an important quality of care issue by the public. According to a national study in 1995 by a consortium of physician-owned insurance carriers, delay in diagnosis of breast cancer is the most common reason that physicians are sued and the second leading cause of indemnity payment (PIAA, 1995). Thus, while there is not direct evidence of process-outcomes links from scientific studies in the diagnostic evaluation of a breast abnormality, there is widespread agreement on the value of ensuring that the appropriate procedures are performed in accordance with the available data and expert consensus. So, in order to monitor the quality of this important aspect of breast cancer care, process measures must be developed which rely primarily on expert opinion and consensus regarding the process-outcomes link.

Treatment of breast cancer, the third phase in the continuum of care, has extensive evidence for process-outcomes links from multiple randomized controlled trials. Surgery has been the primary treatment for localized breast cancer since Halsted popularized the radical mastectomy in 1894 (Halsted, 1894). More recently, randomized controlled trials have demonstrated equivalent survival with a modified radical mastectomy or with breast conserving surgery followed by radiation therapy (Fisher et al., 1985; Sarrazin et al., 1984; Veronesi et al., 1981). In addition, having a choice of surgery appears important to a woman's subsequent quality of life; studies have not demonstrated any difference in overall quality of life between women who received breast conserving surgery and those treated with modified radical mastectomy among patients as long as they were *offered a choice* of primary therapy (Kiebert et al., 1991; Ganz et al., 1992). The most recent NIH Consensus Conference statement for the treatment of early-stage breast cancer specifies breast conserving surgery as the preferred mode of therapy for the majority of women with Stage I and II breast cancer (NIH, 1990).

Strong process-outcomes links also exist for treating women with local or regional breast cancer with chemotherapy or hormone therapy in addition to surgery and radiation. Systemic treatment with chemotherapy or hormone therapy after all identifiable cancer has been removed surgically is termed adjuvant therapy. The goal of adjuvant therapy is to decrease future recurrences thereby improving

survival. However, the issue of adjuvant therapy in breast cancer raises another important consideration: when, in spite of strong evidence in the literature for a process-outcomes link, is the impact on outcome so small that the process should not be considered requisite for quality care? Adjuvant systemic therapy with either chemotherapy or hormone therapy has been demonstrated in randomized controlled clinical trials to improve survival in all women with breast cancer, though the benefit in women with very favorable prognoses is extremely small (EBCT, 1992; 1998). In addition, a small but significant improvement in both overall and disease-free survival is obtained from combined treatment with chemotherapy and Tamoxifen compared with Tamoxifen alone in women with estrogen receptor positive, regardless of the patients' age (Fisher et al., 1997). These studies demonstrate improved breast cancer outcomes with chemotherapy, Tamoxifen, and perhaps both treatments together in all patients. However, given that the absolute benefit is extremely small in patients with a good prognosis (2% improvement in 10 year survival for low risk patients), the benefits may not outweigh the risks of adjuvant treatment in these patients (Osborn et al., 1996). So while most experts would agree that all women with involved lymph nodes, large tumors, and even moderate-size tumors should receive adjuvant therapy, whether it is essential to offer treatment to women with extremely small tumors has not yet been clearly established. Thus, in spite of clear evidence of process-outcomes links for adjuvant therapy in breast cancer, the determination of which patients should be treated remains an issue of expert judgement and consensus.

The extensive clinical research of breast cancer care provides firm grounding for process-outcomes links. Receiving timely and appropriate screening, diagnostic evaluation, and treatment for breast cancer is critical for good outcomes. In addition, the strong process-outcomes links make it possible to rely on process measures in addition to outcomes measures to evaluate the quality of care. A solid foundation of scientific research provides an extremely favorable vantage point from which to begin to evaluate the quality of care in breast cancer.

B. Studies of Breast Cancer Quality

As discussed earlier, to assess quality of care, measures of structure, process, or outcome can be used. If outcome measures are used, care must be taken to account for differences in outcomes that might simply reflect differences in patient selection or case-mix. If structure or process measures are used, they should be associated with the desired outcomes. In addition, in order to make inferences about quality, a structure or process measure must be compared to a standard. Simply comparing variations in the structure or process of care does not provide an evaluation of the quality of care, although it may point to potential quality problems that merit further inquiry. Researchers have compared patterns of breast cancer care in the United States across geographic region, patient age, and race/ethnicity. We will be discussing these studies later in this report. In some cases, these patterns of care studies use explicit outcomes or process measures from which inferences about quality can be made. However, in many cases they simply report variations in care which suggest need for further research using explicit criteria and appropriate case-mix adjustment.

We review the available data on breast cancer processes of care and outcomes. In addition, in a separate section, we summarize selected studies of variations in breast cancer care where these suggest potential quality problems. For this review, we identified all English-language articles in the current MEDLINE database containing one of the following keyword or subject heading along with the key word breast cancer: quality of care, quality and process, quality and outcomes, appropriateness, patterns of care. In addition, we performed searches on each of the subtopics of mammography, biopsy, pathology/cytology, and radiation therapy using the same search strategy. After reviewing all of the abstracts, we selected all studies that report on breast cancer outcomes and processes of care in the United States.

1. Breast Cancer: Process

As described previously in this review, data from randomized controlled trials provide extensive evidence for an association of process with outcomes in breast cancer. Therefore, it is possible to establish treatment standards and develop process measures that compare actual care against the standard. Two process measures appear in the breast cancer literature, whether patients receive radiation therapy after breast conserving surgery and whether patients with locally advanced breast cancer receive adjuvant therapy, both of which are considered requisite for good quality care. These two process measures have been used by researchers to evaluate quality of care in different regions of the United States, with increasing patient age, according to patient, and across hospitals.

Rates of radiation therapy after breast conserving surgery suggest that some parts of the United States may have poor quality breast cancer care. In the nine SEER Registry areas, the percentage of women receiving radiation therapy after breast conserving surgery in 1985-86 was increasing but still varied greatly (Farrow et al., 1992). Though Iowa had the greatest increase in the use of radiation therapy for women with Stage I-II breast cancer treated with breast conserving surgery, it still had the lowest use of all nine areas with only 60% of women in 1985-86 receiving radiation. Seattle had the highest use with 81% in 1985-86 receiving radiation therapy after breast conserving surgery. While some women may refuse radiation therapy, one would hope that an informed discussion of the treatment options would have led many women who did not wish to receive radiation to choose mastectomy as their primary treatment. Thus, one might expect that only a small number of women would opt for breast conserving surgery without radiation.

A limitation of this study, and many of the other studies we will be discussing that also rely on cancer registry data, is that the reliability of the data on treatment collected by the cancer registries has not been systematically evaluated. Thus, the low rates of radiation therapy after breast conserving surgery reported in this study may reflect incomplete data and not poor quality of care.

Alternatively, the low rates of radiation therapy after breast conserving surgery reported by Farrow et al. (1992) may reflect the practice in the 1980's but may not accurately describe the quality of care currently in the United States. Two recently published studies of breast cancer care in selected populations suggest that the quality of breast cancer care may have improved, at least for some women. Hillner et al. (1997) used 1989-1991 data from the Virginia Cancer Registry to evaluate the quality of care of 918 Virginia women ages 64 and younger with Stage I-III breast cancer who had Blue Cross Blue Shield health insurance. In this patient population, 82% of women who underwent breast conserving surgery received radiation therapy. In addition, 83% of women 50 and younger (who were assumed to be premenopausal) with node-positive disease received adjuvant chemotherapy. The authors were unable to assess the use of adjuvant hormone therapy through the claims data.

Guadagnoli et al. (1998) compared the care received by women diagnosed in 1993-1995 with Stage I or II breast cancer at 18 hospitals in Massachusetts with that of women at 30 hospitals in Minnesota diagnosed in 1993. In contrast to the previous studies that relied upon administrative data, these authors collected data about breast cancer treatment from medical records, patient surveys, and physician surveys. While the hospitals participating in Massachusetts were randomly selected (20 were originally selected but two refused to participate), those from Minnesota were part of a consortium formed by the Healthcare Education and Research Foundation, although they include about 60% of patients hospitalized in Minnesota. Among women treated at these institutions from 1993-1995, 84% treated in Massachusetts and 86% treated in Minnesota received radiation therapy after breast conserving surgery. In addition, 97% and 94%, respectively, of pre-menopausal women with node-positive breast cancer received adjuvant chemotherapy. By contrast, only 63% of post-menopausal women in Massachusetts and 59% in Minnesota who had positive lymph nodes and positive estrogen receptor status received adjuvant hormone therapy.

Another study used claims and cancer registry data to examine treatment of local breast cancer in Pennsylvania during 1986-1990, the time period in between the Farrow (1992) study and the Hillner (1997) and Guadagnoli (1998) reports (Young et al., 1996). It found that 82% of women received radiation therapy after breast conserving surgery. Of note, there was substantial variation in the use of radiation therapy by the patient's type of insurance: 45% of Medicaid beneficiaries received radiation therapy, compared with 78% of Blue Cross/Blue Shield subscribers, and 88% of Medicare enrollees.

Osteen et al. (1994) also reported substantial variation in radiation therapy after breast conserving surgery by insurance type according to data voluntarily reported to the American College of Surgeons by 1,011 hospital tumor registries on 24,356 patients diagnosed with breast cancer in 1990, though the percentages differed substantially from those reported by Young et al. (1996). Osteen et al. found that only 52% of Medicare patients and 67% of patients with private insurance received radiation therapy after breast conserving surgery (compared with 88% and 78% in Young et al.) in the national convenience sample. On the other hand, Medicaid patients were more likely to receive radiation therapy in this study, with 61% reported to have had radiation after breast conserving surgery. Among HMO patients, 65% received radiation therapy following breast conserving surgery. The highest rates of radiation therapy in this study were in patients treated at military facitilies where 74% of women had radiation after breast conserving surgery. The differences reported by these two studies may represent regional variations in the quality of care (better care for Medicare and private insurance in Pennsylvania compared with nation as a whole and worse care for Medicaid patients); however, it may simply reflect variation in the completeness of the

different sources of data. The study by Young et al. may have more complete data on radiation therapy since they used claims and registry data while Osteen et al. relied solely on hospital registries for information about radiation therapy.

An earlier study by Johnson et al. (1994) which just looked at the use of adjuvant therapy in women with Stage and I and II breast cancer diagnosed in 1983-1989 also found high adherence to National Cancer Institute Consensus Conference guidelines among cases treated at community hospitals participating in the Community Clinical Oncology Program (CCOP) in 1983-1989, control hospitals in 1983-1986, and a random sample of controls from the SEER registries in 1987-1989. The CCOP was initiated by the NCI in 1983 to increase community participation in clinical research (Kaluzny et al., 1995). In this study, the proportion of women with node-positive breast cancer receiving adjuvant hormone or chemotherapy was highest in 1988 after the NCI released a Clinical Alert advising physicians of the potential benefits in women with *node-negative* disease (the NIH Consensus Conference on adjuvant therapy in breast cancer had only advised adjuvant therapy in women with node-positive disease.) The proportion of women with node-negative breast cancer treated with adjuvant therapy increased from 26% in the quarter before the Clinical Alert to 54% in the quarter following its release. During the same period, the proportion of women with node-positive breast cancer treated with adjuvant therapy increased from 81% to 90%. A year after the Clinical Alert, the percentage of node-negative patients remained elevated above baseline at 46%, while the percentage of node-positive women treated with adjuvant therapy had fallen back to the baseline rate at 79%. It is not known if the Clinical Alert had a transient spill over effect on treatment of node-positive disease. The proportion of women receiving adjuvant therapy in this study approaches a level one would expect if all women who would benefit from adjuvant therapy were being offered treatment.

However, we are limited in our ability to make inferences about the quality of treatment with adjuvant therapy from these data. The authors do not distinguish between the cases that were from CCOP participant hospitals, control hospitals, or the SEER registry so we do not know if the rates of adjuvant therapy varied between these populations. Facilities that chose to participate in the NCI's CCOP are more likely to have an interest in cancer treatment and may have much higher rates of adherence to NCI guidelines than other facilities in the community. These issues notwithstanding, these data, and those reported by Hillner (1997) and Guadagnoli (1998), demonstrate high levels of adherence to treatment standards for adjuvant therapy in breast cancer in selected patient populations, with the notable exception of the low use of hormone therapy in post-menopausal patients.

The higher rates of radiation therapy reported by Hillner (1997), Guadagnoli (1998), and Young (1998) compared with the previous Farrow study (1992) may reflect a general improvement in the quality of breast cancer care in the United States. However, given the selected patient populations in the Hillner (1997), Guadagnoli (1998), and Young (1998) studies, these data must be interpreted cautiously, especially when attempting to generalize from these results to the entire U.S. population. The Hillner (1997) study includes only women younger than 60 with private fee-for-service health insurance. And, while the Guadagnoli (1998) study is not limited to a privately insured population, it is limited to patients treated at hospitals that agreed to participate and therefore may be providing a higher standard of care. The Young (1996) study compared women with Blue Cross/Blue Shield, Medicare, and Medicaid and found that women with Medicaid received much poorer quality care. Nevertheless, these studies suggest that at least some women in the United States had access to high quality breast cancer care by 1995. However, even among these women there is cause for concern since only 60% of postmenopausal women with node-positive estrogen receptor positive cancers received adjuvant hormone therapy.

Other studies using breast cancer process measures suggest that the proportion of women receiving standard treatment decreases with age, though the results of the published literature are not concordant. Several studies using cancer registry data to assess care for Medicare patients have shown that elderly women are not receiving recommended treatments for breast cancer as often as younger women, even when controlling for comorbid illness. Using New Mexico Cancer Registry data from 1984 to 1986, Goodwin et al. (1993) found that while only 43% of women 85 and older, 84% of women 75 to 84 received definitive treatment for localized breast cancer compared with 92% of women 65 to 74 (Goodwin et al., 1993). Definitive breast cancer treatment was defined as lumpectomy or excisional biopsy followed by radiation therapy or mastectomy. Age remained significant even when controlling for women's access to transportation, physical activity levels, income, social support, ability to perform activities of daily living, mental status, and the presence of other medical illnesses. In a comparable population of women in Virginia in 1985-1989, also using cancer registry data, Hillner et al. (1996) found that while the reported

number of women 65 to 69 receiving radiation therapy after breast conserving surgery was inappropriately low at 66%, only 7% of women 85 and older had received radiation therapy (Hillner et al., 1996). In addition, although adjuvant therapy is recommended for all patients with node positive disease, just 44% of patients with positive lymph nodes received any adjuvant therapy, and only 33% received hormone therapy. Several studies, using data on treatment collected by the SEER cancer registries, have also noted that the use of radiation therapy in women who have undergone breast conserving surgery is lower than expected and declines with age (Farrow et al., 1992; Lazovich et al., 1991; Ballard-Barbash et al., 1996). Ballard-Barbash et al. (1996) found that 76% of women 65 to 69 received radiation therapy after breast conserving surgery for Stage I or II cancer, 68% 70 to 74 year olds, 56% of 75 to 79 year olds, and 24% of women 80 years old and older did (Ballard-Barbash et al., 1996). They found that while controlling for differences in comorbidity was associated with a decrease in the frequency of radiation therapy after breast conserving surgery across all age groups, the decline with age persisted.

An earlier study, using data obtained directly from the medical record in seven Southern California hospitals in 1980 through 1982, also found rates of appropriate breast cancer treatment declining with age (Greenfield et al., 1987). Greenfield et al. (1987) reviewed patients' medical records to determine whether they had received diagnostic testing, staging evaluation, and treatment that was consistent with stage-specific consensus recommendations at that time (e.g. radiation therapy after breast conserving surgery, adjuvant chemotherapy for pre-menopausal women with node-positive breast cancer, adjuvant hormone therapy for post-menopausal women with node-positive estrogen receptor positive tumors.) While the proportion of women receiving the recommended diagnostic and staging evaluations did not vary with age, the proportion receiving the recommended treatment did. According to the authors, 83% of women 50 to 69 received the recommended breast cancer treatment compared with 67% of women 70 and older. This difference remained significant when controlling for comorbidity, stage of breast cancer, and hospital where treatment occurred.

In contrast, in a study of postmenopausal women ages 50 and older with Stage I and II breast cancer treated in 1993 at thirty hospitals in Minnesota, using data collected through patient self-report, survey of the treating physician, and the medical record, Guadagnoli et al. (1997) found 92% of women with node positive breast cancer received some form of adjuvant therapy (Guadagnoli et al., 1997). While the likelihood of women with node-positive breast cancer receiving adjuvant therapy did appear to decline slightly with age, this was not statistically significant. The use of adjuvant therapy did decline in women with node-negative breast cancer, and this was true both of chemotherapy and hormone therapy. Seventy-three percent of women 50 to 59 years old with node negative breast cancer received adjuvant therapy compared with 67% of women 60 to 69, 56% of women 70 to 79, and 36% of women 80 and older; however, these age-associated differences were not significant after adjusting for marital status, education, income, health maintenance organization membership, tumor size, lymphatic invasion, estrogen receptor status, grade, type of primary surgery, history of breast and other cancer, and the severity of comorbid disease. Use of adjuvant hormone therapy in node-negative disease also declined in women 80 and older to 34% from only 52% in the other age groups. Since the authors did not stratify on tumor size or estrogen receptor status, we cannot determine if the subset of women with node-negative disease but larger tumors were more likely to receive adjuvant therapy, suggesting that care was being provided in manner consistent with the scientific evidence and medical consensus of the time.

This last study by Guadagnoli et al. (1997) reports appropriately high rates of adjuvant therapy and thus contradicts the other studies which suggest that problems exist with the quality of care provided for elderly women with breast cancer in the United States (Goodwin et al., 1993; Hillner et al., 1996; Farrow et al., 1992; Lazovich et al., 1991; Ballard-Barbash et al., 1996; Greenfield et al., 1987). Of note, the Guadagnoli (1997) study uses multiple data sources, including patient self-report, physician report and the medical record to obtain information about treatment. All of the authors who report poorer adherence to standard treatment in elderly patients with the exception of Greenfield (1987), used cancer registry data, which again raises the issue of the reliability of the cancer registry data on processes of care. Another possible explanation for the higher rates of adjuvant therapy in the elderly in the study by Guadagnoli et al. (1997) is that the quality of care has improved over time; they report data from 1993 whereas the other studies include data from 1980-1986. Perhaps with the dissemination of the results clinical trials performed in the 1980s, the use of adjuvant therapy and of radiation therapy after breast conserving surgery in the elderly has appropriately increased. Alternatively, there may be regional variations in the quality of care that explain the discrepancies in the results of these studies. Minnesota may have better quality care for breast cancer than the rest of the United States. In any case, given the preponderance of data suggesting

that compliance with standard therapy for breast cancer in older women in the United States is low, even in the face of an isolated study which shows excellent quality of care in 30 hospitals in Minnesota, these results highlight potential problems in quality of breast cancer care that warrant further investigation.

We only found one study that attempts to compare the quality of breast care across hospitals, significant differences in process measures do appear to exist. In a study by Hand et al. (1991), the interquartile range (25^{th} percentile to 75^{th} percentile) for hospitals in Illinois not providing radiation therapy after breast conserving surgery was 17% to 75% (Hand et al., 1991). For not providing adjuvant therapy, the interquartile range was 30% to 56%, and for not performing an estrogen receptor test on the pathologic specimen it was 4% to 14%. Based upon these intriguing results, it seems likely that variability in the quality of breast cancer care does exist across hospitals. Further evaluation of quality of care across hospitals, as well as the causes of this variation and ways to improve the quality at all facilities is warranted.

In conclusion, studies of processes of care which have compared rates of radiation therapy after breast conserving surgery and adjuvant therapy for locally advanced breast cancer suggest that problems do exist with the quality of care received by many women with breast cancer in the United States. Many women (perhaps as many as 40%) do not appear to be receiving indicated radiation therapy after breast conserving surgery, and in some regions of the United States, for example Iowa, it appears that an even greater percentage of women are not receiving radiation therapy. In addition, older women are less likely to receive radiation therapy after breast conserving surgery. Rates of radiation therapy after breast conserving surgery also vary across hospitals suggesting that some hospitals are providing poorer quality breast cancer care than others. Equally concerning, many women do not appear to be receiving adjuvant chemotherapy (perhaps as many as 60%.) Also, older women are less likely to receive adjuvant therapy. These findings must be interpreted with caution, however, since most of the data reported is from the 1980s and are based on cancer registry data, the reliability of which is not known.

2. Breast Cancer: Outcomes

Only a few studies have examined breast cancer outcomes. The main outcomes evaluated are stage at diagnosis and 5-year survival. These outcomes have been compared across patients who live in different regions of the United States, patients of different age, patients of different race/ethnicity, and patients with different types of health care coverage.

Only one study, to our knowledge, has compared breast cancer outcomes across regions of the United States. Farrow et al. (1992) used 1983-1991 data from the nine SEER registries to compare five-year survival rates in women with breast cancer; rates ranged from 71.0% in Iowa to 79.9% in Hawaii (Farrow et al., 1992). These differences persisted after adjusting for age and stage. However, other important case-mix adjusters, such as the presence of comorbid illnesses, were not included in their model, so interpretation of these results is difficult. Thus, it is not clear whether these regional variations in survival from breast cancer reflect differences in the patient populations or regional differences in quality of care.

Researchers have also compared breast cancer stage at diagnosis, which can be viewed as an outcome of screening, across different age groups. Although stage of breast cancer at diagnosis is an outcome of early detection with mammography, and a useful measure when comparing comparable populations (e.g. across health plans), any observed differences with age may not reflect differences in the quality of care but instead represent age-related differences in disease biology or the effectiveness of screening. These make it difficult to attribute any age-related differences in stage to differences in the quality of screening unless differences are observed across another category (e.g. age-related differences in stage which vary across health plans, regions of the country, Medicare HMO plans, etc.) The studies examining breast cancer stage and age do not make such comparisons, so we are limited in our ability to draw conclusions on the quality of screening from them.

Several studies have examined the crude relationship between stage and age. In a sample of 508,724 breast cancer cases diagnosed from 1985 to 1993, using data on AJCC stage submitted voluntarily to the American College of Surgeons from hospital cancer registries, Winchester et al. (1996) found that the proportion of early stage cases was lowest in women younger than 35 (Winchester et al., 1996). The ratio of early (defined as stage 0 and I) to late stage cases was 0.5 for women ages 35 and younger, 0.9 for women 35 to 49, and 1.1 for women 50 and older. In contrast, in a sample of breast cancer cases obtained from the New York State Tumor Registry on New York City residents diagnosed with breast cancer between 1980 and 1985, Mandelblatt et al. (1995) found that the proportion of early stage cases (defined

here as in situ or localized disease using the SEER extent of disease staging system) was *lower* in women 50 to 64 than among women 20 to 49 or 65 and older (Mandelblatt et al., 1995). In this study, the odds ratio for having late stage disease was 1.19 for women 50-64 compared with women 20-49. Finally, obtaining staging information from the medical record, Greenfield et al. (1987) did not find a difference in the proportion of AJCC Stage I and II breast cancer cases among women ages 50 to 69 and ages 70 and older diagnosed in seven Southern California hospitals between 1980 and 1982 (Greenfield et al., 1987).

These studies describe variations in stage by age. In spite of focusing on an important outcome measure, stage at diagnosis, conclusions about quality of breast cancer care from these data are limited. First, none of the studies compare variations in the stage of breast cancer with age across another category (e.g. age-related differences in stage which vary across health plans, regions of the country, Medicare HMO plans, etc.) Second, because two of the three studies do not report the actual rates of breast cancer diagnoses by stage, we are unable to compare results across the studies (in order to make inferences about differences by region or over time.) Third, the use of different age cut-offs makes even simple comparisons of the crude results of these studies impossible. Thus, we cannot draw any conclusions about variations in breast cancer screening outcomes according to patient age.

Breast cancer outcomes have also been studied in relation to patient race/ethnicity, although this has been limited to comparisons of African Americans and Caucasians. African American women have been noted to have both worse disease-free and overall survival than do Caucasian women (Osteen et al., 1994). The lower survival of African American women is at least partially attributable to the fact that they appear to be diagnosed with later stage disease (Breen and Figueroa 1996; Polednak et al., 1986; Hillner et al., 1996; Winchester et al., 1996; Hunter et al., 1993; Mandelblatt et al., 1996; McCarthy et al., 1998). In the National Cancer Institute Black/White Cancer Survival Study, a case-control study which included white and black women diagnosed with breast cancer in 1985, the odds of being diagnosed with Stage III or IV breast cancer was 1.68 for black women compared with white women, controlling for age, city of residence, tumor grade, a clinical breast exam by a physician in the previous six years, and a history of patient delay (Hunter et al 1993). Approximately half of the excess risk for Stage III-IV disease among black women was explained by tumor grade, having a clinical breast exam in the previous six years, and a history of patient delay. So, while differences in tumor biology between women of different race/ethnicity may play a role in both stage at diagnosis and stage-specific survival from breast cancer, access to medical care for early detection and timely diagnosis appears extremely important.

In this same study, the risk of death was 2.2 times greater for black women than white women (Eley et al., 1994). However, this difference was no longer statistically significant after controlling for breast cancer stage, tumor pathology, treatment, comorbidity, and sociodemographic characteristics.

In a recent case-control study by Lannin et al. (1998) of women diagnosed with breast cancer at the University Medical Center of Eastern Carolina from 1985-1992, the odds of presenting with late stage disease (Stage III and IV) was three times greater for African American women than for white women (Lannin et al., 1998). When demographic variables were included in a logistic regression model, the odds ratio for late stage disease in African American women decreased to 1.8. When measures of cultural beliefs (e.g. beliefs such as the devil can cause you to get cancer or air causes cancer) were included in the model, African American women no longer had increased odds of late stage disease. Thus, it appears that the worse survival among African American women with breast cancer can be explained in large part by determinants of access to health care.

Indeed, access to mammography appears to account for much, if not all, of the variation in stage at diagnosis of breast cancer attributable to race (Mandelblatt et al., 1995; Breen and Figueroa 1996). In a recent study comparing mammography use between black and white women ages 67 and older, there were no differences in stage at diagnosis among black and white women who were regular mammography users, though the odds of being diagnosed with late stage breast cancer was 2.54 greater for black women compared with white women who had not participated in screening mammography (McCarthy et al., 1998). In addition, studies of women who receive their medical care through the Department of Defense, and would therefore have the same ability to access care, demonstrate no difference in stage between Caucasian, African-American and Hispanic women diagnosed with breast cancer (Zaloznik 1995, 1997). Among women receiving their health care through the Department of Defense in 1980-1992, 49% African-Americans were diagnosed with Stage 0 or I breast cancer compared with 46% of Hispanic women and 57% of Caucasian women (the differences were not statistically significant.) These studies suggest that differences in stage at diagnosis among women of different races may be explained by their relative access to health care. Thus, while African Americans may be receiving poorer quality breast cancer care, it is

difficult to separate the effect of the care received from access when evaluating outcomes alone. Process measures may therefore be more sensitive at identifying differences in the quality of breast cancer care by race. It is worth noting, however, that lack of access in and of itself may be considered an indicator of poor quality.

Breast cancer outcomes appear to be worse for patients with Medicaid and no insurance than for patients with private insurance. By linking data from the New Jersey Cancer Registry with hospital-discharge abstracts on patients diagnosed with breast cancer in 1985-1987, Ayanian et al. (1993) compared the stage, stage-specific survival, and overall survival for breast cancer patients with private insurance, Medicaid, and those without insurance (Ayanian et al., 1993). Approximately half as many uninsured and Medicaid patients were diagnosed with Stage I disease compared with the privately insured patients (30% of privately insured vs. 17% of Medicaid vs. 16% of uninsured women had Stage I breast cancer,) although the demographic characteristics of the three groups varied significantly making the data difficult to interpret. Survival was also worse for the uninsured and Medicaid patients compared with privately insured patients, even when controlling for age and stage at diagnosis. The relative risk of death was 1.49 for uninsured women and 1.40 for Medicaid compared to women with private insurance. In another study, using data from the participating American College of Surgeons hospital registries, later stage diagnoses and poorer survival were once again observed in the uninsured and the Medicaid population, though they did not find significant differences in stage at diagnosis or survival between other types of insurance, including traditional private insurance, HMO coverage, Medicare, military benefits, or CHAMPUS (Osteen et al., 1994). These studies suggest an "enabling" aspect to insurance coverage. That is, stage at diagnosis and mortality are likely to be affected by lack of access to care; however, once access is established, type of insurance does not appear to be particularly influential in the outcome. Or perhaps, the effect of insurance is overwhelmed by the other demographic variables discussed above, which also can covary with insurance.

Only two studies have directly looked at the effect of managed care on breast cancer care. Potosky et al. (1997) compared Medicare patients enrolled in a group model HMO in Seattle and a staff model HMO in San Francisco with fee-for-service Medicare patients in Seattle and San Francisco diagnosed from 1985-1992 using SEER data (Potosky et al., 1997). They found that breast cancer patients in the two HMOs were diagnosed with earlier stage disease compared with women in fee-for-service Medicare; 57% of HMO enrollees in San Francisco and 58% in Seattle versus only 46% of San Francisco women and 51% of Seattle women in fee-for-service were diagnosed with Stage 0 or I breast cancer. In addition, women in the HMO had better overall survival, even when controlling for comorbidity, than women in fee-for-service. The risk ratio for death in ten years for San Francisco HMO enrollees compared with women in fee-for-service was 0.70 (95% confidence interval 0.62-0.79), while in Seattle the risk ratio was 0.86 (95% confidence interval 0.72-1.03), after adjusting for age, race, education, stage, and comorbidity. Thus, outcomes of Medicare patients with breast cancer in these two established HMOs were at least as good as that of Medicare enrollees in fee-for-service, and for enrollees of Kaiser in San Francisco the quality of care appeared better than for enrollees of fee-for-service Medicare.

In contrast, Lee-Feldstein et al. (1994), also using cancer registry data, found significantly worse five-year survival among women with breast cancer treated at HMO hospitals in Orange County, California in 1984-1990 compared with women treated at teaching hospitals, small community hospitals, and large community hospitals. Even after controlling for age, tumor size, number of positive lymph nodes, *and type of treatment* (e.g. breast conserving surgery with radiation vs. no radiation), patients treated at an HMO facility had an increase risk of dying of 1.63 compared with the reference group, small hospitals. In contrast the risk of death at teaching hospitals was not significantly different from small hospitals and it was lower at 0.74 at large hospitals. Interpretation of these results is limited because the effect of the HMO plan cannot be isolated from the effect of the hospital. No definition for "HMO hospital" is provided and the study does not describe whether all the HMO hospitals were part of one or several HMO plans. The type of HMO is also not described (e.g. staff, group, network, or IPA model.) Also, it is not clear if HMO patients could be treated only at HMO hospitals or if they could obtain care at other facilities as well. In addition, this study did not control for comorbidity, socioeconomic status, and insurance status, and length of enrollment in the HMO.

The contradictory results presented by the Potosky and Lee-Feldtstein studies underscore the importance of case-mix adjustment as well as the tremendous heterogeneity in the rapidly evolving managed care systems. Research is needed to understand the impact of these systems on the quality of care.

In summary, while some of these studies do demonstrate some variations in outcomes by region and with race and insurance status; we are limited in our ability to draw any conclusions about quality of care from these data. First, the studies do not use the same definitions of stage nor the same age groupings making comparisons difficult. Second, much of the data presented does not include the crude proportions of cases by stage or the survival data but risk ratios comparing different groups of patients. Third, few of the studies attempt to adjust for differences in case-mix. And, finally, data are from cases diagnosed in the 1980s so we do not have information on the quality of the current breast cancer care in the United States. Nonetheless, the breast cancer outcomes of 5 year survival and stage at diagnosis appear worse for patients who are African American and patients with Medicaid or no health insurance. In both cases, it is likely that decreased access to the health care system plays an important role. Studies that include more recently diagnosed breast cancer cases, with appropriate case-mix adjustment, are needed to evaluate current outcomes and identify important predictors.

C. Variations in Patterns of Care for Breast Cancer: An Issue of Quality?

Some studies examine variations in patterns of care but do not include a benchmark or standard so direct inferences about the quality of care are not possible. Although, these studies can suggest issues for future quality of care studies. We summarize the patterns of care studies that suggest potential quality of care issues that merit further evaluation.

1. Variations in Rates of Breast Conserving Surgery

Many studies have compared the proportion of women who receive breast-conserving surgery instead of mastectomy. The decision about which type of breast surgery to undergo depends upon the size of the primary tumor, the skill and preferences of the surgeon, and the preferences of the patient. The proportion of women who would choose breast-conserving surgery if they were presented with enough information to make an informed choice is not known. Since there is no benchmark for what percentage of women should receive breast-conserving surgery, whether any identified variation in the proportion of women who the more conservative surgery is associated with the quality of care cannot be ascertained. However, widespread differences in the percentage of women who receive the two types of surgery would suggest that some women are not able to completely exercise their choice.

The proportion of women who receive breast-conserving surgery instead of mastectomy as primary surgical treatment for breast cancer varies dramatically in different regions. Studies conducted in the 1980s after the results of the randomized controlled trials comparing the two procedures were published demonstrate widespread variation in the use of breast conserving surgery in the United States. The proportion of all women 65 and older with non-metastatic breast cancer receiving breast conserving surgery ranged from 4% in Kentucky to 21% in Massachusetts in 1986 according to Medicare data (Nattinger et al 1992). Across the nine areas with SEER Registries, Seattle appeared to have the highest rates of breast conserving surgery in 1983-1989 with 34% of women with Stage I and II disease receiving breast conserving surgeries (Farrow et al., 1992; Lazovich et al., 1991; Samet et al., 1994). And, the rates of breast conserving surgery among Medicare patients with non-metastatic disease appear to have increased minimally from 1986 to 1990 (14.1% to 15.0%) (Nattinger et al., 1996). While patient age, the sociodemographic characteristics of the communities, hospital characteristics, and the availability of radiation therapy appear to affect the proportion of women who undergo breast conserving surgery; marked geographic variation in the use of the procedure persists even after adjusting for these characteristics (Farrow et al., 1992; Lazovich et al., 1991; Nattinger et al., 1992; Samet et al., 1994). Although the proportion of women undergoing breast conserving surgery if all eligible women were offered the procedure is not known, one reason for the wide regional variations could be variation in the quality of breast cancer care.

A recent study by Guadagnoli et al. (1998) using data collected from the medical record and a patient survey reported much higher utilization of breast conserving surgery but still substantial regional variation when comparing women treated in 18 hospitals in 1993-1995 in Massachusetts and 30 hospitals in 1993 in Minnesota. Overall 64% of women in Massachusetts and 38% of women in Minnesota with Stage I and II breast cancer received breast conserving surgery. However, upon excluding patients with contraindications to breast conserving surgery (e.g. prior breast conserving surgery, tumor multifocal, tumor centrally located, etc.), the proportions of women receiving breast conserving surgery increased to 74% and 48%, for Massachusetts and Minnesota, respectively. Importantly, 27% of women in Minnesota and 15% of women in Massachusetts who underwent mastectomy, even though they were eligible for

breast conserving surgery, reported that their surgeon had not discussed breast conserving surgery with them. This suggests that a significant proportion of the variation in rates of breast conserving reflects women not being given a choice of procedure.

Data compiled by Wennberg et al. in the Dartmouth Atlas of Health Care using 1992-1993 Medicare claims data show that widespread variation in the proportion of women offered breast conserving surgery remains (Wennberg et al., 1996). Without stratifying on stage in this predominantly 65 and older population, the percentage of inpatient cancer surgery that was a breast conserving procedure ranged from 1.4 to 48.0% by Medicare hospital referral region. In light of the data from Guadagnoli et al. (1998) that women who undergo mastectomy are more likely not to have been provided information about breast conserving surgery, the persistent widespread regional variation in the performance of breast conserving surgery would appear to indicate that many women are not being offered a choice.

Other studies have found significant variation in the use of breast conserving surgery according to hospital characteristics (Nattinger 1992, 1996). A study by Nattinger et al. (1996), found that 55% of breast conserving surgeries performed on Medicare patients occurred in only 10% of the hospitals submitting claims in 1986 to 1990 (Nattinger 1996). Increased use of breast conserving surgery was associated with larger hospital size, the presence of a radiation facility, the presence of a cancer program, having a teaching program, not-for-profit status, and the volume of breast cancer surgeries performed at the hospital. The study by Guadagnoli et al. (1998) also found a positive association between breast conserving surgery and teaching. The odds of breast conserving surgery was 2.4 times for patients treated at teaching hospitals in Massachusetts and 1.5 time in Minnesota compared with patients treated at non-teaching facilities (Guadagnoli et al., 1998). Interestingly, in contrast to the Nattinger study (1996), Guadagnoli et al. found no relation between breast conserving surgery and hospital size or having a radiation facility, perhaps indicating that these factors were no longer significant once higher rates of breast conserving surgery were achieved overall in the community.

These studies suggest that many women are not offered a choice in the type of breast surgery. Further research is needed to determine what proportion of patients are aware that breast conserving surgery is available, if they are given a choice of surgery, as well as which procedure they ultimately received. In order to obtain the information necessary to evaluate this aspect of quality of care, the data will need to be obtained by patient self-report since this level of detail is not contained in cancer registries, claims, or the medical record.

2. Variations in Compliance with American College of Radiology Quality Standards

While many women who undergo breast conserving surgery do not get indicated post-operative radiation therapy, potential quality problems still exist for the women who do get radiation therapy. The 1988 breast cancer Patterns of Care Study performed by the American College of Radiology suggests widespread variation in compliance with standards of quality established by the College, with academic centers demonstrating the greatest compliance, followed by hospital facilities, and free standing facilities having the poorest adherence (Kutchner et al., 1996). For example, immobilization of breast cancer patients receiving radiation therapy, in order to obtain consistent irradiation of the desired target, varied from 80% at academic centers, to 73% in hospital facilities, and only 51% in free standing facilities. Similarly, the use of techniques to decrease divergence of the radiation beam into lung tissue (in order to decrease pulmonary toxicity) ranged from 93% at academic centers, 77% at hospital facilities, and 67% at free standing facilities. In order to understand the relationship between the structure, process, and outcomes of radiation therapy in communities across the country are needed, further research is needed.

3. Variations in the Quality of Breast Diagnostic Procedures

While the process measures described are important for gaining an understanding of the standard of breast cancer care delivered in the United States, simply knowing if a woman received the recommended screening, diagnostic evaluation, and treatment for breast cancer does not provide a complete view of the quality of care. Another component of quality is how well these procedures are performed.

Studies of screening mammography suggest significant variation in both the technical quality of the radiographic images and of the interpretation. A series of evaluations by the Food and Drug Administration (FDA) in 1985, 1988 and 1992 suggested widespread variation in mammography image quality (Houn and Finder, 1997). And, while improvement was noted overall with the percent of acceptable images being 86% in 1992 compared with only 64% in 1985, a significant problem in image quality that was felt to be attributable to differences in technique remained (Segal, 1994). At the urging of

the FDA, American College of Radiology developed a voluntary accrediting process for mammography facilities in 1987 to attempt to correct this problem (Houn and Finder, 1997).

A related issue in the quality of mammography screening is the quality of the interpretation of the image by the radiologist. In a study by Beam et al. (1996), the mammography readings of 108 radiologists from 50 participating facilities randomly sampled from 4611 ACR-accredited mammography centers were compared (Beam et al., 1996). The radiologists' sensitivity for detecting breast cancer (with biopsy as the gold standard) from a screening mammogram ranged from 47% to 100% with a mean of 79%. The specificity also varied tremendously across radiologists with the broad range of 36% to 99%. Concern over the quality of mammography led to Congress passing the Mammography Quality Standards Act in 1992 that establishes standards for physicians interpreting mammograms (Houn and Finder, 1997). The responsibility for its implementation was delegated to the Food and Drug Administration in June 1993. Under this new law, in order for a mammography facility to operate lawfully and be reimbursed by Medicare, it must be certified by the FDA by becoming accredited by an FDA-approved accreditation body.

The quality of the breast biopsy procedure is important so that clinicians making treatment recommendations to women with a breast lesion or with breast cancer have accurate information. Multiple steps in the process of a breast biopsy are critical to ensuring that results are accurate including the biopsy procedure itself (which may be a fine needle aspiration, a stereotactic core biopsy, an open biopsy, or a needle-localization procedure followed by a biopsy,) the tissue preparation, the cytopathology interpretation, assessment of estrogen receptor and progesterone receptor status, and the pathology report which communicates all findings. A few studies have addressed the quality of some of these steps. A single community-based study in New Hampshire assessed the degree of diagnostic agreement among general pathologists reading an investigator-defined set of breast tissue specimens obtained via core and excisional biopsies (Wells et al., 1998). They found overall high agreement among the pathologists for assignment to diagnostic category (kappa coefficient = 0.71) and very high agreement for differentiation of benign versus malignant breast disease (kappa coefficient = 0.95.) Other studies have found that the adequacy of specimens obtained from fine needle aspiration and stereotactic core biopsy vary widely, as do their reported sensitivity and specificity (Stolier 1997; Hayes et al., 1996; Acheson et al., 1997). Yet no published studies have explored the issues that affect the quality of these procedures as performed in the diverse clinical facilities across the United States.

The pathology report is the critical link between pathologist and clinician. Deficits in the pathology report may represent problems with communication or deficiencies in the pathologic evaluation itself. Several studies suggest that variation in the quality of pathology reports for breast cancer specimens exists and warrants further evaluation. In the American Pathologists Q-Probes Study which reviewed 20 breast biopsy specimens and the corresponding pathology report from 434 voluntarily participating surgical pathology laboratories in the United States, Canada, and Australia, only 77% of the pathology reports for the breast cancer cancer cases documented the lesion size (Nakleh et al., 1997). In addition, only 77% contained the margin status (whether the cancer involved the margins of the specimen which is necessary to determine if all of the malignancy was removed by the procedure) and 75% documented whether estrogen and progesterone receptor status had been evaluated. A single institution study of needle localization breast biopsy documented even lower rates of documentation for this important information in the pathology report (Howe et al., 1995). Only 33% of reports in this study commented on the margins of the lesion and estrogen receptor status was found for only 68% of the cases. Such deficits in pathology reports have been the target of quality improvement projects. Hammond and Flinner (1997) reduced the number of incomplete breast cancer pathology reports in a large urban pathology practice in Salt Lake City from 57 of 356 in 1990 to only 2 of 190 in 1995 after instituting a template for the report which included all essential information.

These studies identify multiple steps along the diagnostic evaluation of breast cancer where quality of care may be affected by the quality of the procedure. Poor quality at any step could significant impact the overall quality of care provided. Although having information on every step in the continuum of care would provide us with a comprehensive assessment of the quality of care, generally this is not practical. Acquiring such comprehensive data on quality would be intrusive, time-consuming, and expensive. Nevertheless, considering all of the steps necessary to a multi-step process such as the diagnosis of breast cancer can be extremely valuable in trying to identify the reasons for a quality problem that has been identified (for example, too many women diagnosed with late stage breast cancer.)

D. Breast Cancer: a Focus of Early Cancer Quality Indicators

With the shift in focus in quality assessment from simply exploring variations in patterns of care to accountability, many organizations are developing performance measurement systems in order to be able to evaluate the quality of care delivered at multiple organization levels. As we discussed in the introduction, a number of these organizations have included breast cancer as the first oncologic condition for which they have developed quality indicators. While these measurement systems are in their infancy and will need to be evaluated carefully in order to ensure their validity, these examples provide a perspective on how quality of care for breast cancer, as well as other cancers, will be assessed in the future.

The most recent version HEDIS 3 .0, developed by the National Committee for Quality Assurance has three breast cancer quality indicators, one in the reporting set and two in the testing set (see Table 1) (HEDIS 1997). Health plans are expected to provide information on the measure in the reporting set (proportion of women who receive screening mammography) while measures in the testing set are optional and under evaluation (proportion of early breast cancers and proportion with appropriate follow-up after an abnormal mammogram.) All three HEDIS breast cancer quality indicators target early detection and diagnosis, not on the care received after breast cancer is diagnosed. NCQA's focus on early detection in breast cancer quality assessment appears to be related to their belief that "early detection remains the most effective way of improving the outcomes of breast cancer (McGlynn 1996)." Because the number of incident breast cancer cases in most health plans is too small to make meaningful comparisons of stage at diagnosis across health plans, NCQA has halted further work on the first measure (McGlynn EA, personal communication). Currently, NCQA is working with RAND to develop the specifications for data collection for the abnormal mammogram follow-up indicator to use with administrative data and chart abstraction. They are also testing the reliability and validity of the measure (McGlynn EA, personal communication).

The Joint Commission on Accreditation of Healthcare Organizations (JCAHO) is another organization that has recently included quality indicators for breast cancer care in its performance measurement set (see Table 1) (JCAHO 1998). However, not all organizations seeking JCAHO accreditation will be required to report on their breast cancer care since JCAHO requires organizations to select several from a menu of performance measures and report data only on those that they choose. Therefore, it is unlikely that these new proposed performance measures will allow us to compare the quality of breast cancer care across facilities in the near future.

As we discussed in the introduction, FACCT has also selected breast cancer as its first oncologic condition to evaluate (FACCT 1997). Breast cancer was selected because of its prevalence, the impact of early-detection on long-term survival, the importance of patient choice in the selection of local treatment (breast conserving surgery with radiation vs. mastectomy,) and the potential effects of the illness on a woman's productivity and quality of life. FACCT's breast cancer quality measures are the most comprehensive of any of the organizations which currently have breast cancer performance measures and include indicators of the process of care, patient satisfaction, and outcomes. These measures are currently being field tested to determine if it is feasible to use them.

The Health Care Finance Administration (HCFA), in collaboration with RAND, is also working on developing tools for breast cancer quality assessment (Katherine Kahn and Marge Pearson, RAND, personal communication, 1998). The RAND research team is reviewing different methodological options for measuring the quality of breast cancer care as part of HCFA's Outcomes Project. To support HCFA's commitment to measuring clinical performance, they are investigating the options for accurately identifying breast cancer patients, articulating multiple domains of breast processes and outcomes of care, collecting and summarizing the existing evidence supporting performance measures in each of these domains, and identifying multiple sources of performance data.

E. Challenges in Breast Cancer Quality Assessment

While more quality of care research has taken place in breast cancer than any other oncologic condition, interpretation of these findings has several limitations. These limitations highlight the challenges of quality of care research in breast cancer, and in many instances, these same challenges apply to the evaluation of the quality of care for other cancers as well. First, there is tremendous publication delay of quality of care research in breast cancer; though most of the data presented in this review is from recently published studies, they report on breast cancer cases diagnosed in the 1980s. Second, while cancer registries are a valuable source of data on cancer care, the reliability of their data on processes of care is not known. Third, in order to compare outcomes, differences in case-mixes need to be adjusted, but obtaining the data to effectively control for these differences in patient populations is often difficult and expensive.

Understanding and meeting these challenges is necessary to be able to effectively evaluate the quality of breast cancer care.

First, there is tremendous publication delay of quality of care research in breast cancer. Most of the quality of care studies cited in this review have publication dates of 1992-1997. However, the data included in these studies is from patients diagnosed with breast cancer in 1980-1990. This is true not only of studies that include the outcome of 5-year survival, which necessarily involves many years before the data is available, but also of those studies that include process measures. The reasons for this delay are clear; however, one cause maybe the delay in the availability of cancer registry data. There is a significant time lag from the time a case is diagnosed until the data are made available to researchers by the cancer registries. Data from the SEER cancer registries are not available for approximately two years after diagnosis (SEER 1998). State cancer registries do not have uniform case reporting requirements. For example, data on cancer cases in California, which has a statewide registry, is not available until October of the second year following diagnosis (data on the 1997 cases will not be available until October 1999) (William Wright, Director, California Cancer Registry, Department of Health Services, California, personal communication, 1997). Such delays in the availability of data are a serious challenge in an era where the health care system is extremely dynamic and accountability is an important aspect of quality evaluation

Second, most of the studies of breast cancer quality have relied upon the data collected by cancer registries. However, while cancer registries have been shown to have approximately 98% reliability in the reporting of cancer cases, the reliability of the data on processes of care collected by cancer registries has not been validated (Zippin et al., 1995). Some of the variability in the results of studies of the relationship of patient age to process of breast cancer care cited earlier may represent variation in the quality of the data. For example, obtaining data directly from the medical record, Guadagnoli et al. (1997) found that 92% of women with node positive breast cancer in Minnesota in 1993 received some form of adjuvant therapy (Guadagnoli et al., 1997). In contrast, using data from the Virginia cancer registry, Hillner et al. (1996) found that only 44% of patients with positive lymph nodes received adjuvant therapy (Hillner et al., 1996). In order to use the data collected by cancer registries to make valid assessment of quality of care, the reliability of registry data on processes of care must be determined.

Third, valid methods are needed to adjust for differences in case-mix. As we have discussed previously, in order to compare outcomes, differences in case-mix need to be adjusted. In some cases, case-mix differences need to be accounted for when comparing process measures as well. Process measures may need to be case-mix adjusted, for example, if the treatment would be harmful or not indicated in a patient with other medical illnesses. Many of the studies we have summarized attempt to control for some differences in case-mix by adjusting for stage of diagnosis, age, sociodemographic characteristics such as income or education, and insurance status. In some cases, they also include a variable for comorbidity; although the validity of the construct used in many cases is not known. A comorbidity index has been developed and validated for breast cancer that can be used with either medical record or claims data (Charlson et al., 1987; D'Houre et al., 1996; Newschaffer et al., 1997). Obtaining such data is often time-consuming and expensive, and may not always be possible. Including data on comorbidity in the information abstracted from the medical record by the cancer registries, using a validated approach, would make this important information available for case-mix adjustment in quality assessment.

F. Summary of Breast Cancer

Studies of breast cancer quality of care have compared process and outcome measures to a standard across regions of United States, patient age, and different hospitals. Rates of radiation therapy after breast conserving surgery and adjuvant therapy for locally advanced disease are lower than expected and suggest that problems may exist with the quality of care received by many women with breast cancer in the United States. Many women do not appear to be receiving indicated radiation therapy after breast conserving surgery, and in some regions of the United States, for example Iowa, it appears that the percentage of women who do not receive radiation is greater. In addition, older women are less likely to receive radiation therapy after breast conserving surgery. Rates of radiation therapy after breast conserving surgery also vary across hospitals suggesting that some hospitals are providing poorer quality breast cancer care than others. Equally concerning, many women do not appear to be receiving adjuvant chemotherapy. Although these findings must be interpreted with caution since much of the data are from breast cancer cases diagnosed in the 1980s and primarily collected by cancer registries, they suggest serious problems exist with quality of care provided to women with breast cancer in the United States. We need to evaluate

the quality of care, using reliable process and outcome measures, of a national sample of newly diagnosed breast cancer patients to determine the quality of *current* practice.

VII. CHILDHOOD CANCER CASE STUDY

Childhood cancer is rare. About 11,000 children less than 20 years old are diagnosed with cancer each year in the U.S. (Bleyer, 1995). Over the past several decades, great advances have been made in the treatment of many childhood cancers. Forty years ago, the mortality rate for childhood cancer was about 80%; now it is about 20% (Bleyer, 1997).

We chose childhood cancer as one of our case studies for three principal reasons. First, children differ from adults in three clinically important ways: their developmental trajectory; their dependency on parents and other caregivers; and the type, intensity, and prevalence of conditions that affect them (Jameson and Wehr 1993). These differences are relevant to cancer as well as to general health issues. Consequently, quality assessment for children is different and in many ways more complex than quality assessment for adults. For example, because childhood is a period of rapid change, it is necessary to develop quality measures that are age- and developmentally-appropriate. An indicator looking at correct response to an abnormal blood pressure reading would have to take into account the different definitions of normal as the child grows and matures.

Second, there has generally been much less work on quality of care for children than for adults, so we wanted to determine whether this was true in the cancer field as well.

Third, most children with cancer receive their care as part of a clinical trial (i.e., they are on a protocol), and we expected that would influence the quality of care they receive. Almost 95% of U.S. children with cancer less than 15 years old are registered by one of the two national collaborative groups—the Children's Cancer Group (CCG) and the Pediatric Oncology Group (POG)—and about 70% of children are enrolled in at least one clinical trial. By comparison, only about 2% of adults with cancer participate in NCI-sponsored clinical trials (Bleyer, 1997). We also anticipated that—regardless of what work has already been done to measure quality—the high participation rate in protocols and the current level of centralization through CCG and POG would facilitate opportunities to measure quality in the future.

In this section, we will describe several aspects of quality of care as well as work that could be done in the future. Our review turned up very few studies relevant to quality of pediatric cancer care. Therefore, many of the studies we report do not provide direct evidence of quality of care but instead provide an indication of what methodological groundwork exists for conducting future work in this area.

A. Process Measures

We were not able to identify studies reporting on how well care is being delivered to children with cancer. However, one potential marker for quality of care is the level of missing data in multi-site studies in which individual sites sent a series of data elements to the registry (e.g., Kisker et al., 1980; Gaynon et al., 1987). When there are missing data, we do not know if the clinicians have not done what they need to do to produce the data (e.g., ask a question on history, perform a part of the physical examination, or order a particular laboratory test), or if they simply have not reported the data element to the central registry. A record abstraction at the actual site would elucidate the meaning of missing data in such studies.

For example, Gaynon et al. (1987) described the use of a computerized system for tracking data from multiple sites participating in pediatric cancer protocols. The study found that data on key indicators (e.g., no FAB [French-American-British] classification, no report of bone marrow assessment 7 days after starting therapy) were not reported for many children. It is not clear whether providers were providing the care but not reporting it to the central office, or whether the missing data represent laboratory tests that were never performed and care that was never delivered. If the data are missing because the care was not given, this would serve as evidence of compromised quality. The paper also reported on whether doses of chemotherapeutic agents were correct (within 10% of the recommended dose). If the patient received a dose other than the specified dose and no justification for the change was documented, questions would be raised about the quality of care provided. Because this study did not collect information on reasons for giving doses that did not follow the protocol, we cannot comment on the quality of care. However, the system described in the article lays the groundwork for what could be a useful tool for measuring quality of care.

Although quality of pediatric cancer care assessment is in its infancy, there is a strong infrastructure on which to build. Because most children with cancer are treated as part of a protocol, data

on them and their care are already being reported to central registries. This data collection process could serve as the foundation on which quality assessment could be built.

B. Biological Outcomes: Mortality

We were able to identify only a few studies that specifically looked at outcomes for children. Several studies from the 1970s and 1980s used outcome measures to compare quality between referral centers and community-based centers delivering cancer care. While these studies might not be considered traditional quality of care studies, they serve as a useful illustration of what can be done in pediatric cancer and what work might be done in the future.

Shah (1986) compared cancer care at a community center and an academic center before and after efforts were made to improve care at the community center. Initially, relapse free survival was higher at the academic center. After the intervention, the rates were similar at both centers. The study stratified by risk category of ALL (Acute Lymphoblastic Leukemia), but it did not risk adjust for factors that might have influenced both the site parents chose for their children's care and the survival of the child, such as parent vigilance in managing the child's illness. (The changes instituted to improve care at the community center were substantial and could have influenced parents deciding where to take their child. The type of parents who might ensure that their child complies with regimens and who might provide strong support for their child might have initially preferred the academic center but might have been more willing to go to the community site after it was upgraded—in other words, there could have been a change in the types of parents and children at the two sites.) Despite possible limitations to the study design, the improved outcome was probably attributable to the changes made at the community center. Another approach, though, would have been to compare the processes of care. This approach might have provided more conclusive evidence that the differences in outcomes were attributable to the quality of care provided; it also would have helped identify what both centers could have done to improve care.

Other studies from the U.S. and Great Britain comparing outcomes between community and referral centers had similar findings, although they were also limited in their ability to risk adjust or control for factors that would have led some patients to be referred and not others (Duffner et al., 1982; Kramer et al., 1984; Meadows et al., 1983; Stiller, 1988). It should be noted, though, that one study from the same period found similar outcomes between a community-based setting and pediatric hematologist/oncologists when the former were collaborating with the latter (Kisker et al. 1980; Strayer et al., 1980), and there have been subsequent efforts to provide better coordination of care between community-based sites and cancer centers.

C. Potential Outcome Measures for Childhood Cancer

While survival is generally considered the most important outcome in childhood cancer, other outcome measures may also be of value in evaluating quality of care. Recent development and validation of tools to assess functional status, pain control, and patient satisfaction allow them to be further studied as potential outcome measures for quality assessment in childhood cancers.

1. Functional Status

Outside of the cancer field, there have been only a few measures of functional status developed for children, in part because of the challenge of developing measures that are age-appropriate. The Functional Status II(R) measure is designed for use with children with chronic physical conditions and includes items on physical mobility, behavior, and social functioning (Stein and Jessop, 1990). The Child Health Questionnaire measures the physical and psychosocial well-being of children five years and older (Landgraf, 1996), and the Child Health and Illness Profile-Adolescent Edition measures health status of adolescents ages 11 through 17 years old (Starfield et al., 1995).

There have been various efforts to create functional status measures specifically for children with cancer and for survivors of childhood cancer. As far as we know, none of these measures has been used for quality assessment. However, we report on them here because some could be used or adapted for quality assessment purposes.

Lanksy et al., (1979) modified the Karnofsky scale (described above) for use with children (Lansky et al., 1979). Lansky subsequently developed a single dimension play-performance scale as a general functional measure for children with cancer (Lansky et al., 1987).

Olson et al. (1993) combined eight existing measures to look at social competence and emotional health from the perspective of the child, parent, and teacher. Each of the eight components takes between 5

and 20 minutes to complete, so the full instruments are unlikely to be used routinely in the clinical setting specifically for quality assessment purposes. However, if they were administered for clinical purposes, their findings could be incorporated into a quality assessment scheme.

A British study developed the multi-attribute system for classifying the health status of survivors of childhood cancer (Feeny et al., 1992). The system includes sensation (health, speech, and vision), mobility, emotion, cognition, self-care, pain, and fertility. The system is based on functional capacity rather than performance; in other words, it measures the extent to which deficits in health status affect functioning rather than the level at which an individual chooses to function. The ability to function is measured from poor to good or normal, and the respondent is asked to compare the person's capacities to those capacities that would be considered developmentally appropriate for age. For example, the instrument asks whether the subject learns and remembers schoolwork normally for age. Researchers found the system easy to use for classifying a sample of long-term survivors of high-risk ALL (Feeny et al., 1993). Although the system was not designed to assess quality, it could be adapted for such use, especially if it were applied to children who are currently receiving care or who recently completed care.

Other work is underway in the area of quality of life measurement. The American Cancer Society Workshop on Quality of Life in Children's Cancer: Implications for Practice and Research, in 1995, produced an extensive framework for addressing quality of life issues for children with cancer in a clinical setting. This framework can serve as a resource for researchers working on developing specific quality of life measures (Lauria et al., 1996). Jenney et al. (1995) and Eiser and Jenney (1996) also provide reviews of quality of life measures used in pediatric oncology.

Functional status and quality of life measures could be considered for use in assessing quality of care for children during and after therapy. For example, one might compare functional status of children at several institutions as one measure of quality of care. It might be the case that children treated at a hospital with a strong case management and psychological support program had higher functional status, or a hospital that provided better treatment or a hospital with a strong post-cancer follow-up program might have better functional status outcomes. Of course, it would be important to risk-adjust, and it would have to be clear that the health care system had an influence over these outcomes. Another challenge would be the need for a large enough sample size to make valid comparisons and to provide data in a timely enough fashion for use in quality comparisons.

2. Pain

An important issue in cancer treatment is whether patients receive adequate pain control. There have been various efforts to measure pain both during and after treatment. Adequate pain management would be a reasonable measure to include as part of quality assessment. It is certainly important for any pain measure used for quality assessment to be both valid and reliable. Clinically useful and valid measures have been developed for evaluating intensity of pain in children 3 years old and older (McGrath et al., 1990). These could be adapted if necessary for use in quality measurement. Pain can also be assessed in younger children, although it typically requires inference from behavioral and physiological responses before children have developed more advanced language skills (McGrath et al., 1990). Adapting such measures for quality assessment purposes might be more challenging.

3. Satisfaction and Patient Ratings

Satisfaction and patient ratings in pediatrics typically involve assessment of parents' satisfaction and ratings. However, there has been some effort to find out what children think, especially older children and adolescents.

There has been limited work done on patient ratings in pediatric cancer. The CCG recently published a survey to assess the informed consent process for pediatric cancer trials. This survey includes items asking parents about the consent process as well as their satisfaction with the process. A convenience sample of 23 parents who piloted the survey had a mean satisfaction score of 17, on a scale from 7 to 21, where 21 indicated highest satisfaction (Kodish et al., 1998).

D. Summary of Childhood Cancer

In summary, there has been only limited work done in assessment of quality of cancer care for children. However, the epidemiology of cancer, the diagnostic strategies, and the recommended treatments are different for children compared to adults. Therefore, it is important that efforts to develop measures and systems to assess quality of cancer care consider these differences. The relatively small number of

children with cancer will make quality assessment challenging. There may not be enough children with the same condition at most centers to allow for simple comparisons. One solution is to develop generic measures that allow assessment across multiple cancers, treatment strategies, and ages. Another is to develop a scoring system that will allow aggregation of scoring across multiple quality measures covering multiple types of cancer and types of care (Schuster et al., 1997).

Although there has been little work done in quality of pediatric oncology, there is a registry infrastructure in place and a tradition of reporting information that will facilitate quality assessment. Pediatric oncologists are used to reporting patient data to a central registry. Furthermore, most children are cared for in academic or referral centers or at community sites associated with such centers. The centralization of care will also make data collection and reporting easier.

VIII. PROSTATE CANCER

A. Burden of Disease

Adenocarcinoma of the prostate (prostate cancer) is the most common form of cancer diagnosed in men. In 1998, more than 185,000 cases are expected to be diagnosed, and more than 39,000 men will die from this disease (American Cancer Society Surveillance Research, 1998). The number of cases has varied considerably in recent years as a result of adoption of prostate-specific antigen (PSA) testing (Potosky et al., 1995). In 1993, before PSA testing became widespread, 165,000 new cases of prostate cancer were identified in the United States. With the August 1994 FDA approval of PSA testing, it became possible to diagnose and treat tumors before they became palpable. As a result, the number of detected cases has grown substantially. This rise largely reflects an increase in early detection rather than an increase in incidence. And, in fact, as the early cases of prostate in the population are being identified and treated, the incidence of prostate cancer is reverting to pre-PSA levels. So, while 334,500 new diagnoses were expected for 1997 (Parker, et al., 1997) this estimate was adjusted downward, to fewer than 210,000 before the end of that same year (Wingo, et al., 1997).The risk of prostate cancer increases with age, with the average age of diagnosis about 65 years old and the median age about 72; it is relatively rare in men younger than 50 years old. Five-year survival rates are also very high. Because of the age distribution and the survival rates, prostate cancer ranks 21st in years of life lost due to cancer (regardless of gender).

Assessing quality of care for prostate cancer detection and treatment is especially difficult. In breast cancer, we have scientific evidence and professional consensus regarding the efficacy of screening for certain groups of women. We also have evidence from randomized trials and other studies that document efficacy of early treatment, and the comparative efficacy of alternative treatments. In contrast, for prostate cancer, definitive evidence supporting the efficacy of early detection and early treatment are pending the results of two clinical trials. NCI's Prostate, Lung, Colorectal, and Ovarian Cancer (PLCO) screening trial is examining the efficacy of early detection of prostate and other cancers, and the Prostate Cancer Intervention Versus Observation Trial (PIVOT) is a randomized trial that is addressing the efficacy of primary treatment of early stage prostate cancer by surgery as compared with conservative management.

Yet, as we are awaiting the results of these studies, nearly two hundred thousand men will be diagnosed and will seek treatment for prostate cancer each year. As we discussed in the introduction, to evaluate quality of care using process measures requires comparison against a standard. Yet, without essential information about the efficacy of early detection, we do not have a standard against which to assess the quality of screening or primary treatment. Measuring outcomes also has limitations in prostate cancer since the illness often progresses very slowly so long follow-up times are needed to show differences in survival or disease-progression. Differences in rates of recurrence or survival may be a result of treatment, but if treatment is not proven to be efficacious, then observed differences in outcomes across providers may simply reflect differences in patient selection. Without essential information about the efficacy of early detection and primary treatment, we are limited to examining whether adverse effects of treatment (complications) can be minimized.

Below we discuss issues in the early detection of prostate cancer and its treatment and review what is known about possible links between the structure and process of prostate cancer treatment and its results. Presently, we have some limited understanding that there can be variations in the results of treatment from facility to facility, but we know very little about the reasons for these variations. The challenge for prostate cancer health services researchers is to begin to define what processes are essential to delivering

efficacious and effective care for prostate cancer patients so that we can lay a groundwork to assess the quality of care received by these patients.

B. Early Detection: Current Recommendations for Prostate Cancer Screening

Methods used to detect prostate cancer include digital rectal examination (DRE); prostate-specific antigen (PSA); and imaging, most often using transrectal ultrasound (TRUS). However, currently the research literature does not provide evidence that screening reduces mortality from prostate cancer, and there is no uniform professional consensus regarding whether routine screening for prostate cancer should be performed.

The broadest screening recommendation is offered by The American Cancer Society (ACS). ACS recommends that annual PSA testing be offered to men over age 50 who have at least a 10-year life expectancy. Certain high risk groups (those with a family history including two or more first degree relatives and African American men) are counseled to begin testing at earlier ages, e.g. age 45 (von Eschenbach et al., 1997).

The American Urological Association recommends offering the PSA test to men over the age of 50 who present for evaluation of prostatic disease symptoms after counseling the patient on the risks and benefits of the test. For those at high risk-- that is, patients with a family history or African American patients-- the recommended age is 40 (Correa, 1998).

National Cancer Institute's PDQ documents that there is insufficient evidence to establish whether improvements in survival are associated with prostate cancer screening by DRE, TRUS, or PSA testing. The US Preventive Services Task Force (1996) does not recommend routine screening for prostate cancer because of insufficient evidence regarding efficacy; it recommends that men who request screening be given information about the risks and benefits of early detection and treatment. Similarly, the Veterans' Administration guidelines for preventive screening include a recommendation that men over age 50 receive a discussion about the risks and benefits of prostate cancer screening including PSA testing; however, no specific recommendation for routine screening is indicated (Wilson and Kizer, 1998).

In the absence of clear evidence or uniform recommendations for prostate cancer screening, we cannot use screening rates as a quality indicator. The results of NCI's Prostate, Lung, Colorectal, and Ovarian Cancer (PLCO) screening trial may better inform recommendations for routine prostate cancer screening. Further, improvements in the specificity of the PSA test may make the use of serum marker testing more effective in screening for prostate cancer.

C. Primary Treatment

1. Pre-treatment Work-up

Clinical staging is done using all information available prior to primary treatment, including DRE, imaging, and biopsy results (American Joint Committee on Cancer, 1997). Clinical staging of prostate cancer may be reported using one of two systems: modified American staging or TNM (tumor-node-metastasis; American Joint Committee on Cancer) staging. Current methods for clinical staging may result in a substantial proportion of cases being understaged (up to 59%), with a smaller proportion being overstaged (about 5%) (Bostwick et al., 1994).

Numerous studies have illustrated the prognostic usefulness of pre-treatment PSA, clinical stage, and Gleason score in predicting post-treatment outcomes such as risk of recurrence (American Joint Committee on Cancer, 1997; Pisansky et al.,1997). Partin et al. (1993) have developed tabulated estimates for risk of extracapsular extension using Gleason score (indicating tumor differentiation) and PSA. The goal of this approach is to attempt to improve prediction of pathological stage for patient counseling and treatment planning. Estimates from their original tables have been improved by pooling data from patients across multiple facilities (Partin et al., 1997), but concerns about patient representativeness may limit the use of these tables as decision aids for physicians. The ratio of free to total PSA may improve the accuracy of staging (Pannek et al., 1998). Catalona (1996) suggests that free PSA may correlate with the potential aggressiveness of localized prostate cancer. However, Pannek et al. (1996) found that it did not provide additional utility in predicting pathologic stage after controlling for Gleason score and clinical stage for early stage prostate cancer patients.

In addition to PSA, Gleason score, and stage, patient co-morbidity can provide independent prognostic information about treatment outcomes. Two of the authors of this paper (McGuigan, Reifel) recently completed interviews with prostate cancer treatment experts in urology and radiation oncology at academic treatment centers around the U.S. We found good agreement across prostate cancer specialists

about the importance of co-morbidity assessment as part of the pre-treatment work-up, but there was considerable variation in the methods used for such assessment. Suggested information to use included: Karnofsky Performance Status, patient self-reported activity levels, obesity, and history of cardiac disease, vascular disease, pulmonary disease, hypertension, diabetes, and surgeries. Pre-treatment urinary, bowel, and sexual functioning were most commonly assessed by patients' verbal reports. Some physicians reported using the AUA symptom score to assess obstruction; formal assessment of potency, voiding symptoms or continence were rarely performed on a routine basis.

While there is evidence in the literature that PSA, stage, Gleason score, and patient comorbidity provide useful prognostic information when treating patients, there is no evidence whether performing these assessment prior to initiating treatment improves patient outcomes. Given the absence of process-outcomes links for the pretreatment evaluation, developing process measures for this aspect of the prostate cancer care would have to be based completely on expert opinion. And, at present, there are no specific guidelines for the staging work-up or pretreatment assessment of patient co-morbidity.

2. Choice of Treatment Modality

The modality used for primary treatment of prostate cancer varies depending on the stage of disease, age or life expectancy, and patient preference.

Treatment of localized prostate cancer (T1 or T2) can include surgery (radical prostatectomy), radiation therapy (external beam, brachytherapy, or conformal radiation therapy), or expectant management (watchful waiting). However, surgical treatment is not recommended for patients whose life expectancy is less than 10 years. In addition, conformal radiation therapy is still being studied for efficacy and side-effects (as compared to standard external beam radiation therapy,) but has not yet been widely adopted as standard practice among radiation oncologists. We do not have definitive evidence about the comparative efficacy of alternative treatment modalities for treating early stage prostate cancer. The information used to make such decisions may have varying accuracy depending on its source: for example, clinicians' assessments of post-treatment complications have been found to greatly underestimate the rate reported by patients themselves (Litwin et al., 1998). This finding may suggests that a potentially important area for quality assessment is differences across providers in the patient counseling process.

A specific recommendation from the American Urological Association's clinical guidelines on the management of clinically localized prostate cancer is that all alternative treatment modalities (radical prostatectomy, radiation therapy—external beam, interstitial treatment—and expectant management) should be presented to every patient (Middleton, 1995). Thus, a potential quality indicator could include assessment of whether these recommendations were followed by urologists.

Treatment for advanced prostate cancer is palliative. Data from randomized controlled trials demonstrate a survival benefit as well relief from bone pain with treatment with androgen ablation, which can include orchiectomy alone, monotherapy with an luteinizing hormone-releasing hormone (LHRH) analogue or "maximal androgen blockade" with either orchiectomy or an LHRH analogue and antiandrogen therapy (Garnick, 1996). When prostate cancer progresses on androgen ablation therapy, treatment is less effective; however, ketoconazole or aminoglutethamide, estramustine, suramin, mitoxantrone with prednisone or steroids alone can improve pain control can quality of life (Garnick, 1996.) Because patients with prostate cancer that has metastasized to the bone often suffer from excruciating pain, a primary focus in the care of patients with metastatic prostate cancer is the control of their pain, either with narcotics, radiation therapy, or chemotherapy. So, while advanced prostate cancer is not curable, multiple treatment options exist and there is evidence in the scientific literature that they improve quality of life and, in some case, prolong survival. Thus, there is sufficient evidence for process-outcomes links in advanced prostate cancer that process measures could be developed to evaluate the quality of care.

D. Quality of Care Research in Prostate Cancer

We searched MEDLINE for any studies with data on structure, process or outcomes of prostate cancer. Given the lack of evidence for definitive process-outcomes links, it is not surprising that there is very limited information on the quality of care of men with prostate cancer.

From the case series presented in the prostate cancer clinical literature, we know that rates of complications can vary substantially. Such information is illustrated graphically in the American Urological Association's summary report on the management of clinically localized prostate cancer (Middleton et al., 1995). Point estimates of complications resulting from primary treatment of prostate

cancer vary widely across facilities even when stratifying by treatment modality: surgery, external beam radiation, or brachytherapy (interstitial radiation treatment or seed implants). Following radical prostatectomy, rates of stress incontinence range from less than 10% to 50%, and impotence rates range from 25% to 100% across series reports. Complications following external beam randiation included proctitis, with rates ranging from less than 10% to over 50%; cystitis, ranging from 0% and 80%; and impotence, ranging from less than 10% to nearly 40%. Similarly, complication rates reported for brachytherapy range from 0% to 75% for proctitis, less than 10% to 90% for cystitis, and from less than 10% to 75% for impotence.

While these widely varying complication rates may reflect differences in quality of care, it is difficult to draw conclusions based on this type of information (Wasson et al., 1993). First, there may be differences in the way that the data was collected which could account for the variations in complications rates. Second there may be systematic diffferences in patient case-mix (disease severity and co-morbidities) across facilities, and these differences need to be accounted for before comparing outcomes across institutions. Even if series reports could be case-mix adjusted, there is usually little information available to link differences in results to differences in the technical process of care. Finally, these series reports data from only a small number of providers and these are often large, academic clinics. The results seen for such providers may not represent those of other institutions or clinics.

An early Patterns of Care study examined the association between types of radiation equipment (a structure measure) and localized prostate cancer treatment outcomes (Hanks et al., 1985). Facilities that used cobalt units were found to have higher stage-adjusted rates of disease recurrence than facilities that used linear accelerators or betatron. Use of cobalt equipment was also correlated with other structural indicators: these facilities had lower percentages of patients who were staged, had lower staff/patient ratios, and were more likely to have part-time therapists as compared with national averages. From these results, the authors recommended that facilities using cobalt units should upgrade their treatment equipment, treat palliative patients only, or close. While this study helped to show that cobalt units were not as good, the evidence available when the care was provided did not indicate that cobalt was inappropriate. Therefore, this study was quite valuable in showing how to improve care, but it is not evidence of poor quality. Quality needs to be judged by the level of knowledge and standards in place at the time the care was delivered.

A new study by Lu-Yao et al (under review) looked at the relationship between outcomes and the number (or volume) of patients receiving surgical treatment for prostate cancer from a surgeon or facility, a structure measure. Using Medicare claims data, they found that high volume facilities had significantly better surgical outcomes than facilities treating fewer patients. High volume facilities had more favorable rates of survival, complications and readmission following treatment by radical prostatectomy than lower volume facilities. High volume facilities also showed shorter lengths of stay. The results suggest a dose-response relationship, where facilities in the highest volume quartile showed the best outcomes, followed by those in the third quartile, etc. These analyses controlled for differences across facilities in patient age, race, year of surgery and hospital teaching status.

E. Information Sources About Quality of Prostate Cancer Care

There are a number of initiatives in place that will help us to understand better what structural characteristics or technical processes of care may be associated with better outcomes for treating prostate cancer.

An excellent example is the American College of Radiology's (ACR) Patterns of Care Study (Kramer and Herring, 1976). Begun in 1971, this study has collected information periodically from a national sample of radiation oncology facilities. As summarized previously, an early Patterns of Care study examined the association between types of radiation equipment and localized prostate cancer treatment outcomes (Hanks et al., 1985). Patterns of Care data have been collected in 1972-74, 1977, 1979, 1984, 1989 and 1994. Data collected at each of these waves have included structure and process information; some years have also included patient outcome measures. For example, 1989 assessed only structure and process measures, while 1994, which has not yet been made publicly available, included outcome assessments as well.

Medicare claims can also provide valuable information which can be used to study variations in patterns of care, and if linked with the appropriate case-mix information may also be a source of data on quality of care. As we discussed earlier, Lu-Yao has used Medicare data to study variations in prostate

cancer outcomes following radical prostatectomy across facilities and in relation to the volume of patients treated.

The NCDB provides information about the treatment of patients in community hospitals as well as at academic medical centers (e.g. Mettlin et al., 1997a, b). These data have been supplemented with surveys of cancer registries to report on the results of prostate cancer treatment.

The National Comprehensive Cancer Network has developed an outcomes database that pools information across NCI-designated Comprehensive Cancer Center participants. In 1997, NCCN began the creation of a uniform outcomes reporting system (Weeks, 1997); results of this effort are pending.

The American Urological Association is developing a Documented Outcomes Collection System (DOCS) to assess patient outcomes for a selected number of conditions. It may collect information related to prostate cancer outcomes in the future.

F. Need for quality measures

Presently there are no quality indicators or performance measures for prostate cancer treatment assessed by HEDIS, FAACT, or JCAHO.

Given the large number of men who are diagnosed and treated for prostate cancer each year, it is important to be able to measure and report on the quality of care available for these patients. We need to develop reliable and valid quality indicators for prostate cancer, and we need to be able to report results for a nationally representative sample or census of treatment providers at the facility or provider level. This information needs to be made available to prostate cancer patients and their physicians as they consider options for where and how to be treated.

G. Summary of Prostate Cancer

Prostate cancer provides a particular challenge for quality of care assessment: we have methods for early detection, but we do not yet have definitive information about the efficacy of early detection. We have a number of treatment modalities for early stage disease, but we do not yet have definitive information about the efficacy of early treatment. Primary treatment itself can have complications that may be relatively short-term and manageable, but for many patients, treatment of prostate cancer can result in long-term problems such as urinary or bowel incontinence or impotence.

One candidate indicator for prostate cancer quality assessment is to identify whether information about alternative treatment modalities was presented to patients, as recommended by the American Urological Association's practice guidelines (Middleton, et al., 1995). A second candidate indicator may be to assess rates of surgical treatment among men with life expectancies less than 10 years, using age 70 to represent a proxy for 10 year remaining life expectancy for the average patient (a high rate of surgical treatment in men with a life-expectancy less than 10 years would be an indicator of *poor* quality.).

We presently have limited empirical information about what structures and technical processes of care are associated with better outcomes for men treated for prostate cancer, and data sources can differ widely in the types of information they report. However, there are a number of data systems in place, and initiatives underway, that may substantially improve the quality of information available regarding the treatment of prostate cancer over the next several years. With this information we can begin to move forward to assess the quality of care delivered to these patients.

Information from the Patterns of Care study may be able to provide additional insights into which structures and technical processes of care may lead to better patient outcomes. Similar work should be undertaken to assess the processes and outcomes of surgical treatment using a nationally representative sample of providers.

Two of the authors (McGuigan, Reifel) have been developing candidate quality indicators based upon a structured review of the literature, key informant interviews of prostate cancer experts, and focus groups with patients as part of a RAND study funded by the Bing Foundation. The next phase of this research will be to test the reliability and validity of these candidate measures in evaluating the quality of care, as well as exploring potential process and outcomes links.

The results of the PLCO trial will better inform the decision of whether routine screening for prostate cancer should be performed and for whom; and the results of the PIVOT trial will inform the efficacy of primary treatment of localized prostate cancer by surgery.

Urologists, radiation oncologists, medical oncologists, and cancer health services researchers should become integrally involved in the development of performance measures and indicators for quality of prostate cancer treatment. There should be a greater emphasis on translating information about the

process of primary treatment into information that can be used by patients and their physicians when making decisions about where and how to be treated. Most importantly, the information that is presently reported in the medical literature should be made available to patients and other lay audiences in formats that are useful for informed treatment decision-making.

IX. SUMMARY OF CASE STUDIES

Although much of clinical practice for all three types of cancer has not been studied, breast cancer and childhood cancer have more research supporting clinical care than prostate cancer has (Table 2). The prevalence of childhood cancers is generally much lower than it is for breast and prostate cancer, which makes it more difficult to make statistically valid comparisons. Therefore, breast cancer lends itself best to quality assessment, and perhaps not surprisingly, there has been more work done in assessment of the quality of breast cancer care than of care for the other two types of cancer.

X. SUMMARY

Efforts to measure quality of cancer care in the United States are in the early stages. National organizations conducting quality assessment have primarily focused on prevention and screening, although they are now moving more towards assessing quality of diagnosis and management. Research studies on quality of cancer care have been limited, with most research concentrating on cancers with comparatively higher prevalence and more evidence supporting clinical practice (e.g., breast cancer). We found little use of process measures, although they have been used some for cancer prevention and screening. Research studies have primarily compared outcomes across entities without tying the outcomes to specific practices or risk adjusting to correct for factors that might influence outcomes. Outcomes can be challenging to use because they often have a long time horizon, effect sizes for many interventions are small, risk adjustment is necessary, and it can be difficult to determine which entities or clinicians should be held accountable for the quality of care indicated by outcomes.

Cancer is not the only area that would benefit from more work in the area of quality assessment. The discipline is young, and most types of subspecialty care, whether oncology, pulmonology, or otolaryngology, are in the early stages of addressing quality issues.

There are also challenges facing efforts to measure quality of cancer care. Much cancer care (as well as much health care in general) is not based on clear evidence. There is often a lack of consensus about what constitutes high quality care, and sample sizes of patients at individual sites are often smaller than one would typically want for measuring quality. Nonetheless, these challenges are not unique to cancer, and they do not prevent development and implementation of effective and valuable quality assessment tools for cancer.

Unlike many other subspecialties, the field of cancer care has an elaborate infrastructure in place that could serve as the foundation for comprehensive quality assessment efforts. Cancer specialists have a tradition of conducting organized research with collaboration across multiple sites, and they routinely transfer data to a central registry. Even when patients are not enrolled in a research study, physicians nevertheless report data to cancer registries. Clinicians in oncology have a strong professional identity with active professional organizations, and patients and their families have a number of active and vocal organizations representing their interests. These groups are able to contribute to the development of quality measures for cancer, a system to apply these measures, and a mechanism for disseminating the findings. The field has the opportunity to make great advances in quality assessment that could be used to improve the quality of cancer care that people receive.

Table 1: Quality of Care Indicators Related to Cancer From National Organizations Conducting Quality Assessment

National Committee for Quality Assurance's (NCQA) Health Plan Employer Data and Information Set, Version 3.0 (HEDIS 3.0)
The percentage of adults smokers or recent quitters who received advice to quit smoking from a health professional in the plan
The percentage of women ages 52-69 who had at least one mammogram during the past two years
The percentage of women ages 21-64 who had at least once Pap test during the past three years

The percentage of adolescents who were counseled on substance abuse (including tobacco) during the past year
The percentage of adults in the plan who smoke
The percentage of adult smokers in the plan who quit smoking in the past year
The percentage of adults ages 55 and older who have been screened for colorectal cancer
The percentage with an abnormal Pap smear who received timely follow-up evaluation
The percentage of women with abnormal mammograms who received appropriate follow-up care within 60 days
Satisfaction of women treated for breast cancer with aspects of care they received
Foundation for Accountability (FACCT) Quality Measures-Breast Cancer
The proportion of women age 52-69 who have had a mammogram within a two-year period
The proportion of patients whose breast cancer was detected at Stage 0 or Stage 1
The proportion of Stage I and Stage II patients who indicate that they had adequate information about their radiation treatment options before deciding about treatment
The proportion of breast conserving surgery patients who receive radiation treatment after breast conserving surgery
The mean score for patients' level of satisfaction with breast cancer care including the technical quality, interpersonal and communication skills of their cancer doctor, their involvement in treatment decisions and the timeliness of getting information and services
The mean score for patients on CARES-SF survey which assesses patients' quality of life and experience in living with breast cancer
The probability of disease-free survival for a group of patients, Stages I-IV, who were diagnosed during prior five years
Joint Commission on Accreditation of Hospital Organizations' (JCAHO) Indicator Measurement System (IMSystem)
The proportion of patients undergoing resection for primary cancer of the lung, colon/rectum, or female breast for whom a surgical pathology consultation report is present in the medical record
The proportion of patients undergoing resection for primary cancer of the lung, colon/rectum, or female breast with stage of tumor designated by a managing physician
The proportion of female patients with Stage I or greater primary breast cancer who, after initial biopsy or resection, have estrogen receptor analysis results in the medical record
The proportion of patients with non-small cell primary lung cancer undergoing thoracotomy with complete surgical resection of tumor
The proportion of patients undergoing resection for primary cancer of the colon/rectum whose preoperative evaluation by a managing physician included examination of the entire colon

Sources: NCQA (1997), FACCT (1997), JCAHO (1998)

Table 2: Summary of Findings from Case Studies

	Research Base	**Sample Size**	**Quality Assessment**
Breast Cancer	Good	Large	Some
Childhood Cancers	Good	Small	Limited
Prostate Cancer	Limited	Large	Limited

References

Acheson MB, Patton RG, Howisey RL, Lane RF, Morgan A. Histologic correlation of image-guided core biopsy with excisional biopsy of nonpalapable breast lesions. *Archives of Surgery*. 1997;132:815-821.

Aharony L, Strasser S. Patient satisfaction: what we know about and what we still need to explore. *Medical Care Review*. 1993;50:49-79.

Ayanian JZ, Kohler BA, Abe T, Epstein AM. The relation between health insurance coverage and clinical outcomes among women with breast cancer. *New England Journal of Medicine*. 1993;329:326-331.

Ballard-Barbash R, Potosky AL, Harlan LC, Nayfield SG, Kessler LG. Factors associated with surgical and radiation therapy for early stage breast cancer in older women. *Journal of the National Cancer Institute*. 1996;88:716-726.

Bartlett E, Grayson M, Barker R, Levine D, Golden A, Libber S. An evaluation of generic screens for poor quality of hospital care on a general medicine service. *Medical Care*. 1993;31:394-402.

Beam CA, Layde PM, Sullivan DC. Variability in the interpretation of screening mammograms by US radiologists. Findings from a national sample. *Archives of Internal Medicine.* 1996;156:209-213.

Bernstein S, McGlynn E, Siu A, et al. The appropriateness of hysterectomy, a comparison of care in seven health plans. *Journal of the American Medical Association*. 1993;269:2398-2402.

Berwick D. Continuous improvement as an ideal in health care. *The New England Journal of Medicine*. 1989;320:53-56.

Bleyer WA. The U.S. Pediatric Cancer Clinical Trials programmes: international implications and the way forward. *European Journal of Cancer*. 1997;33:1439-1447.

Bleyer WA. The past and future of cancer in the young. *Pediatric Dentistry*. 1995;17:285-290.

Bloom HJ, Wallace EN, Henk JM. The treatment and prognosis of medulloblastoma in children. A study of 82 verified cases. *American Journal of Roentgenology, Radium Therapy and Nuclear Medicine*. 1969;105:43-62.

Bostwick DG, Myers RP, Oesterling JE. Staging of prostate cancer. *Seminars in Surgical Oncology* 1994;10:60-72.

Breen N, Figueroa JB. Stage of breast and cervical cancer diagnosis in disadvantaged neighborhoods: a prevention policy perspective. *American Journal of Preventive Medicine.* 1996;12:319-326.

Brook RH. Doing more good than harm: the evaluation of health care interventions. In: Warren K, Mosteller F, eds. *Using Scientific Information to Improve Quality of Health Care*. New York: Annals of the New York Academy of Science; 1993:74-78.

Brook R, Chassin M, Fink A. A method for detailed assessment of the appropriateness of medical technologies. *International Journal of Technology Assessment in Health Care*. 1986;2:53-63.

Brook R, Park R, Chassin M, Kosecoff J, Keesey J, Solomon D. Carotid endarterectomy for elderly patients: predicting complications. *Annals of Internal Medicine*. 1990;113:747-753.

Catalona WJ. Clinical utility of measurements of free and total prostate-specific antigen (PSA): a review. *Prostate.* 1996;7(Supple):64-69.

Cella DF, Bonomi AE. Measuring QOL: 1995 update. *Oncology*. 1995;9(Supple):47-60.

Cella DF, Tulsky DS, Gray G, et al. The Functional Assessment of Cancer Therapy (FACT) Scale: development and validation of the general version. *Journal of Clinical Oncology*. 1993;11:570-579.

Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *Journal of Chronic Diseases*. 1987;40:373-383.

Cleary P, McNeil B. Patient satisfaction as an indicator of quality care. *Inquiry*. 1988;25:25-36.

Cohen MR, Anderson RW, Attilio, RM, Green L, Muller RJ, Pruemer JM. Preventing medication errors in cancer chemotherapy. *American Journal of Health-System Pharmacy*. 1996;53:737-746.

Coscarelli Shag CA, Heinrich RL. Development of a comprehensive quality of life measurement tool: CARES. *Oncology*. 1990;4:135-138.

D'Hoore W, Bouchaert A, Tilquin C. Practical consideration on the use of the Charlson comorbidity index with administrative databases. *Journal of Clinical Epidemiology*. 1996;49:1429-1433.

Davies A, Ware J. Involving consumers in quality of care assessment. *Health Affairs*. 1988;7:33-48.

Deyo R, Inui T. Dropouts and broken appointments, A literature review and agenda for future research. *Medical Care*. 1980;18:1146-1157.

Donabedian A. *Explorations in Quality Assessment and Monitoring, Volume 1: The Definition of Quality and Approaches to Its Assessment*. Ann Arbor, MI: Health Administration Press; 1980.

Duffner PK, Cohen ME, Flannery JT. Referral patterns of childhood brain tumors in the state of Connectitcut. *Cancer*. 1982;50:1636-1640.

Early Breast Cancer Trialists' Collaborative Group. Systemic treatment of early breast cancer by hormonal, cytotoxic, or immune therapy. *Lancet*. 1992;399:1-71.

Early Breast Cancer Trialists' Collaborative Group. Tamoxifen for early breast cancer: an overview of the randomised trials. *Lancet*. 1998;351:1451-67.

Eastman P. NCI adopts new mammography screening guidelines for women. *Journal of the National Cancer Institute*. 1997;89:538-539.

Eiser C, Jenney MEM. Measuring symptomatic benefit and quality of life in paediatric oncology. *British Journal of Cancer*. 1996;73:1313-1316.

Eley JW, Hill HA, Chen VW, Austin DF, et al. Racial differences in survival from breast cancer. Results from the National Cancer Institute Black/White Cancer Survival Study. *Journal of the American Medical Association* 1994;272:947-954.

Etheredge L. The structure and accountability for Medicare health plans: government, the market, and professionalism. In: Jones S, Lewin M, eds. *Improving the Medicare Market Adding Choices and Protections*. Washington, D.C.: National Academy Press; 1996:137-158.

FACCT. FACCT Quality Measures. Breast Cancer Measurement Specifications. The Foundation for Accountability, Portland, Oregon, 1997.

Farrow DC, Hunt WC, and Samet JM. Geographic variation in the treatment of localized breast cancer. *New England Journal of Medicine* 1992;326:1097-1101.

Feeny D, Leiper A, Barr RD, et al. The comprehensive assessment of health status in survivors of childhood cancer: application of high-risk acute lymphoblastic leukamia. *British Journal of Cancer*. 1993;67:1047-1052.

Feeny D, Furlong W, Barr RD, Torrance GW, Rosenbaum P, Weitzman S. A comprehensive multiattribute system for classifying the health status of survivors of childhood cancer. *Journal of Clinical Oncology*. 1992;10:923-928.

Feuer EF, Wun LM, Boring CC, et al. The lifetime risk of developing breast cancer. *Journal of the National Cancer Institute*. 1993;85:892-897.

Field M, Lohr K, eds. *Clinical Practice Guidelines: Directions for a New Program*. Washington, D.C.: National Academy Press; 1990.

Fisher B, Bauer M, Margolese R, et al. Five-year results of a randomized clinical trial comparing total mastectomy and segmental mastectomy with or without radiation in the treatment of breast cancer. *New England Journal of Medicine.* 1985;312:665-73.

Fisher B, Dignam J, Wolmark N, et al. Tamoxifen and chemotherapy for lymph node-negative, estrogen receptor-positive breast cancer. *Journal of the National Cancer Institute.* 1997;89:1673-1682.

Fisher RI, Gaynor ER, Dahlberg S, et al. Comparison of a standard regimen(CHOP) with three itensive chemotherapy regimens for advanced non-Hodgkin's lymphoma. *The New England Journal of Medicine*. 1993;328:1002-1006.

Fowler FJ Jr, Barry MJ, Lu-Yao G, Wasson JH, Bin L. Outcomes of external-beam radiation therapy for prostate cancer: a study of Medicare beneficiaries in three surveillance, epidemiology, and end results areas. *Journal of Clinical Oncology.* 1996;14:2258-2265.

Fowler JE Jr, Terrell F. Survival in blacks and whites after treatment for localized prostate cancer. *Journal of Urology.* 1996;156:133-136.

Ganz PA, Coscarelli Shag CA, Lee JJ, Sim M-S. The CARES: a generic measure of health-related quality of life for patients with cancer. *Quality of Life Research*. 1992;1:19-29.

Ganz PA, Schag CAC, Lee JJ, Polinsky ML, Tan S-J. Breast conservation versus mastectomy: is there a difference in psychological adjustment of quality of life in the year after surgery? *Cancer.* 1992;69:1729-1738.

Garnick MB. Prostate cancer: screening, diagnosis, and management. *Annals of Internal Medicine* 1996; 118(10): 804-818.

Gaynon PS, Baum ES, Samsel C, Sather H. Microcomputer-assisted data management for multiinstitutional pediatric clinical cancer trials. *The American Journal of Pediatric Hematology/Oncology*. 1987;9:33-41.

Goodwin JS, Hunt WC, Samet JM. Determinants of cancer therapy in elderly patients. *Cancer.* 1993;72:594-601.

Greenfield S, Blanco DM, Elashoff RM, Ganz PA. Patterns of care related to age of breast cancer patients. *Journal of the American Medical Association.* 1987;257:2766-2770.

Grumbach K, Anderson G, Luft H, Roos L, Brook R. Regionalization of cardiac surgery in the United States and Canada: outcomes, choice, and geographic access. *Journal of the American Medical Association*. 1995;274:1282-1288.

Guadagnoli E, Shapiro C, Gurwitz JH, Silliman RA, et al. Age-related patterns of care: evidence against ageism in the treatment of early-stage breast cancer. *Journal of Clinical Oncology*. 1997;6:2338-2344.

Guadagnoli E, Shapiro CL, Weeks JC, Gurwitz JH, Borbas C, and Soumerai SB. The quality of care for treatment of early stage breast carcinoma. Is it consistent with national guidelines? *Cancer*. 1998;83:302-9.

Guadagnoli E, Weeks JC, Shapiro CL, Gurwitz JH, Borbas C, and Soumerai SB. Use of breast-conserving surgery for treatment of stage I and stage II breast cancer. *Journal of Clinical Oncology*. 1998;16:101-106.

Halsted W. The results of operations for cure of cancer of the breast performed at Johns Hopkins Hospital. *Johns Hopkins Hospital Bulletin.* 1894;4:497.

Hammond EH, Flinner RL. Clinically relevant breast cancer reporting. Using process measures to improve anatomic pathology reporting. *Archives of Patholology and Laboratory Medicine.* 1997;121:1171-1175.

Hand R, Sener St, Imperato J, Chmiel JS, et al. Hospital variables associated with quality of care for breast cancer patients. *Journal of the American Medical Association* 1991;66:3429-3432.

Hanks GE, Diamond JJ, Kramer S. The need for complex technology in radiation oncology. Correlations of facility characteristics and structure with outcome. *Cancer.* 1985;55(9 Suppl):2198-2201.

Hannan E, Racz M, Ryan T, et al. Coronary angioplasty volume-outcome relationships for hospitals and cardiologists. *Journal of the American Medical Association*. 1997;277:892-898.

Hayes MK, De Bruhl ND, Hirschowitz S, Kimme-Smith C, Bassett LW. Mammographically guided fine-needle aspiration cytology of the breast: reducing the rate of insufficient specimens. *American Journal of Roentgenology.* 1996;167:381-384.

Hayward R, Bernard A, Rosevear J, Anderson J, McMahon L. An evaluation of generic screens for poor quality of hospital care on a general medicine service. *Medical Care*. 1993;31:394-402.

Hilborne L, Leape L, Berstein S, et al. he appropriateness of use of coronary artery bypass graft surgery in New York state. *Journal of the American Medical Association*. 1993;269:753-760.

Hillner BE, McDonal KM, Penberthy L, Desch CE, Smith TJ, Patricia Maddux, Glasheen WP, and Retchin SM. Measuring standards of care for early breast cancer in an insured population. *Journal of Clinical Oncology*. 1997; 15:1401-1408.

Hillner BE, Penberthy L, Desch CE, McDonald MK, Smith TJ, Retchin SM. Variation in staging and treatment of local and regional breast cancer in the elderly. *Breast Cancer Research and Treatment.* 1996;40:75-86.

Houn F, Finder C. The Mammography Quality Standards Act and interventional mammography procedures. *The Bulletin of the American College of Surgeons*. 1997;82:14-22.

Howe JR, Monsees B, Destouet J, Seib J, Dehner LP, Kraybill WG. Needle localization breast biopsy: a model for multidisciplinary quality assurance. *Journal of Surgical Oncology.* 1995;58:223-239.

Hsieh MO, Kagle JD. Understanding patient satisfaction and dissatisfaction with health care. *Health and Social Work*. 1991;16:281-290.

Hunter CP, Redmond CK, Chen VW, Austin DF, et al. Breast cancer: factors associated with stage at diagnosis in black and white women. *Journal of the National Cancer Institute*. 1993;85:1129-1137.

JCAHO Website, http://www.jcaho.org, accessed 6/29/98.

Iezzoni L. An introduction to risk adjustment. *American Journal of Medical Quality*. 1996;11:S8-S11.

Jenney MEM, Lane RL, Lurie N. Developing a measure of health outcomes in survivors of childhood cancer: a review of the issues. *Medical and Pediatric Oncology*. 1995;24:145-153.

Johnson TP, Ford L, Warnecke RB, Nayfield SG, Kaluzny A, Cutter G, Gillings D, Sondik E, and Ozer H. Effect of a National Cancer Institute Clinical Alert on breast cancer practice patterns. *Journal of Clinical Oncology*. 1994; 12:1783-1788.

Kahan K, Bernstein S, Leape L, et al. Measuring the necessity of medical procedures. *Medical Care*. 1994;32:357-365.

Kahn K, Rogers W, Rubenstein L, et al. Measuring quality of care with explicit process criteria before and after implementation of the DRG-based prospecive payment system. *Journal of the American Medical Association*. 1990;264:1980-1983.

Kaluzny AD, Warnecke R, Lacey LM, Johnson T, Gillings D, and Ozer H. Using a community clinical trials network for treatment, prevention, and control research: assuring access to state-of-the-at cancer care. *Cancer Investigation*. 1995; 13: 517-525.

Karnofsky DA, Burchenal JH. The Clinical evaluation of chemotherapeutic agents in cancer. In: MacLeod CM, ed. *Evaluation of Chemotherapeutic Agents*. New York, NY: Columbia University Press;1949:191-205.

Kaufman Z, Shpitz B, Shapiro M, Rona R, Lew S, Dinbar A. Triple approach in the diagnosis of dominant breast masses: combined physical examination, mammography, and fine-needle aspiration. *Journal of Surgical Oncology*. 1994;56:254-257.

Kerlikowske K; Grady D; Rubin SM; Sandrock C; Ernster VL. Efficacy of screening mammography. A meta-analysis. *Journal of the American Medical Association*. 1995;273:149-154.

Kiebert GD, de Haes JCJM, van de Velde CJH. The impact of breast-conserving treatment and mastectomy on the quality of life of early-stage breast cancer patients: A review. *Journal of Clinical Oncology*. 1991;9:1059-1070.

Kisker CT, Strayer F, Wong K, Clarke W, Strauss R, Tannous R, Janco R, Spevak J. Health outcomes of a community-based therapy program for children with cancer--a shared-management approach. *Pediatrics*. 1980;66:900-906.

Kitahata M, Koepsell T, Deyo R, Maxwell C, Dodge W, Wagner E. Physicians' experience with the Acquired Immunodeficiency Syndrome as a factor in patients' survival. *The New England Journal of Medicine*. 1996;334:701-706.

Kodish ED, Pentz RD, Noll RB, Ruccione K, Buckley J, Lange BJ. Informed consent in the Childrens Cancer Group. Results of preliminary research. *Cancer*. 1998;82:2467-81.

Kramer S, Meadows AT, Pastore G, et al. Influence of place of treatment on diagnosis, treatment and survival in three pediatric solid tumors. *Journal of Clinical Oncology*. 1984;2:917-923.

Kravitz R, Laouri M, Kahan J, et al. Validity of criteria used for detecting underuse of coronary revascularization. *Journal of the American Medical Association*. 1995;274:632-638.

Kritchevsky S, Simmons B. Continous quality improvement; concepts and applications for physician care. *Journal of the American Medical Association*. 1991;266:1817-1823.

Krook JE, Moertel CG, Gunderson LL, et al. Effective surgical adjuvant therapy for high-risk rectal carcinoma. *The New England Journal of Medicine*. 1991;324:709-715.

Kuban DA, El-Mahdi AM, Schelhammer PF. PSA for outcome prediction and posttreatment evaluation following radiation for prostate cancer: Do we know how to use it? *Seminars in Radiation Oncology*. 1998;8:72-80.

Kutchner GJ, Smith AR, Fowble BL, et al. Treatment planning for primary breast cancer: a patterns of care study. *International Journal of Radiation Oncology, Biology, Physics.* 1996;36:731-737.

Kupelian, PA Kupelian VA, Witte JS, Macklis R, Klein EA. Family history of prostate cancer in patients with localized prostate cancer: An independent predictor of treatment outcome. *Journal of Clinical Oncology.* 1997;5:1478-80.

Landgraf JM, Abetz L., Ware JE. *The CHQ User's Manual*. Boston, MA: The Health Institute, New England Medical Center;1996.

Landis SH, Murray T, Bolden S, Wingo PA. Cancer Statistics 1998. *CA: A Cancer Journal for Clinicians.* 1998;48:6-29.

Lannin DR, Mathews HF, Mitchell J, Swanson MS, Swanson FH, and Edwards MS. Influence of socioeconomic and cultural factors on racial differences in late-stage presentation of breast cancer. *Journal of the American Medical Association.* 1998; 279:1801-1807.

Lansky SB, Cairns NU, Clark GM, Lowman J, Miller L, Trueworthy R. Childhood cancer. Nonmedical costs of the illness. *Cancer*. 1979;43:403-408.

Lansky SB, List MA, Lansky LL, Ritter-Sterr C, Miller DR. The measurement of performance in childhood cancer patients. *Cancer*. 1987;60:1651-1656.

Laouri M, Kravitz R, Berstein S, et al. Underuse of coronary angiography: application of a clinical method. *International Journal for Quality in Health Care*. 1997;9:5-22.

Lauria MM, Hockenberry-EAton M, Pawletko TM, Mauer AM. Psychosocial protocol for childhood cancer. *Cancer*. 1996;78:1345-1356.

Lazovich D, White E, Thomas DB, Moe RE. Underutilization of breast-conseving surgery and radiation therapy among women with Stage I or II breast cancer. *Journal of the American Medical Association.* 1991;266:3433-3438.

Leape L, Bates D, Cullen D, et al. Systems analysis of adverse drug events. *Journal of the American Medical Association*. 1995;274:35-43.

Leape L, Hilborne L, Park R, et al. The appropriateness of use of coronary artery bypass graft surgery in New York state. *Journal of the American Medical Association*. 1993;269:753-760.

Lee-Feldstein D, Anton-Culver H, Feldstein PJ. Treatment differences and other prognostic factors related to breast cancer survival. *Journal of the American Medical Association*. 1994;271:1163-1168.

Levine MC, Guyatt GH, Gent M, et al. Quality of life in stage II breast cancer: an instrument for clinical trials. *Journal of Clinical Oncology*. 1988;6:1798-1810

Litwin MS, Hays RD, Fink A, Ganz PA, Leake B, Leach GE, Brook RH. Quality-of-life outcomes in men treated for localized prostate cancer. *Journal of the American Medical Association*. 1995;273:129-135.

Litwin MS, Lubeck DP, Henning JM, Carroll PR. Differences in urologist and patient assessments of health related quality of life in men with prostate cancer: results of the CaPSURE database. *Journal of Urology*.1998;159:1988-92.

Lu-Yao GL, Potosky AL, Albertsen PC, Wasson JH, Barry MJ, Wennberg JE. Follow-up prostate cancer treatments after radical prostatectomy: a population-based study. *Journal of the National Cancer Institute.* 1996;88:3-4 166-73.

Lu-Yao G, Yao S. Relationships between surgical volume, outcomes and length of stay- A national study of patients undergoing radical prostatectomy. Paper presented at the American Society for Clinical Oncology Conference, Los Angeles CA. 1998; under review.

Lohr K, ed. Medicare: *A Strategy for Quality Assurance*. Washington, D.C.: National Academy Press; 1990:441.

Lohr K, Brook R, Kamberg C, et al. Use of medical care in the Rand health insurance experiment, diagnosis-and service-specific analyses in a randomized controlled trial. *Medical Care*. 1986;24:S1-S87.

Luft H, Garnick D, Mark D, McPhee S. *Hospital Volume, Physician Volume, and Patient Outcomes: Assessing the Evidence*. Ann Arbor: Health Administration Press; 1990:406.

Mandelblatt J, Andrews H, Kao R, Wallace R, Kerner J. Impact of access and social context on breast cancer stage at diagnosis. *Journal of Health Care for the Poor and Underserved.* 1995;6:342-351.

Mark BA, Burleson DL. Measurement of patient outcomes: data availability and consistency across hospitals. *Journal of Nursing Administration*. 1995;25:52-59.

McCarthy EP, Burns R, Coughlin SS, Freund KM, et al. Mammography use helps to explain differences in breast cancer stage at diagnosis between older black and white women. *Annals of Internal Medicine.* 1998;128:729-736.

McGlynn EA. Choosing chronic disease measures for HEDIS: conseptual framework and review of seven clinical areas. *Managed Care Quarterly*. 1996;4:54-77.

McGlynn EA, Senior Analyst, RAND, Santa Monica, California. Personal Communication, 1998.

McGrath PJ, Beyer J, Cleeland C, Eland J, McGrath PA, Portenoy R. Report of the Subcommittee on assessment and methodologic issues in the management of pain in childhood cancer. *Pediatrics*. 1990;86(Supplement):814-817.

Meadows AT, Kramer S, Hopson R, Lustbader E, Jarrett P, Evans AE. Survival in childhood acute lymphocytic leukemia: effect of protocol and place of treatment. *Cancer Investigation*. 1983;1:49-55.

Menck HR, Cunningham MP, Milburn Jessup J. The growth and maturation of the National Cancer Data Base. *Cancer*. 1997;80:2296-2304.

Metlin C, Smart CR: Breast cancer detection guidelines for women aged 40 to 49 years: Rationale for the American Cancer Society reaffirmation of recommendations. *Ca - A Cancer Journal for Clinicians*. 1994;44:248-255.

Mettlin CJ, Murphy GP, Cunningham MP, Menck HR. The National Cancer Data Base report on race, age, and region variations in prostate cancer treatment. *Cancer*. 1997;80:1261-1266.

Mettlin CJ, Murphy GP, Sylvester J, McKee RF, Morrow M, Winchester DP. Results of hospital cancer registry surveys by the American College of Surgeons: outcomes of prostate cancer treatment by radical prostatectomy. *Cancer*. 1997;80:1875-1881.

Middleton RG, Thompson IM, Austenfeld MS, Cooner WH, Correa RJ, Gibbons RP, Miller HC et al. Prostate cancer clinical guidelines panel summary report on the management of clinically localized prostate cancer. *Journal of Urology*. 1995;154: 2144-48.

Miller AB, Baines C, To T, et al. Canadian national breast screening study: . Breast cancer detection and death rates among women aged 40 to 49 years. *Canadian Medical Association Journal*. 1992;147:1459-1476.

Moul JW, Douglas TH, McCarthy WF, McLeod DG. Black race is an adverse prognostic factor for prostate cancer recurrence following radical prostatectomy in an equal access health care setting. *Journal of Urology*. 1996;155:1667-1673.

Nakleh RE, Jones B, Zarbo RJ. Mammogrpahically directed breast biopsies. A college of American Pathologists Q-Probes Study of clinical physician expectations and of specimen handling and reporting characteristics in 434 institutions. *Archives of Pathology and Laboratory Medicine* 1997; 121:11-18.

National Committee for Quality Assurance. *HEDIS 3.0, Volume 1: Narrative: What's in It and Why It Matters*. Washington, D.C.: National Committee for Quality Assurance; 1997.

National Institutes of Health. *Adjuvant Therapy for Patients with Colon and Rectum Cancer. NIH Consensus Statement*. Bethesda, MD: National Institutes of Health. 1990;8:1-25.

National Institutes of Health. *NIH Consensus Statement. Treatment of Early Stage Breast Cancer*. Bethesda, MD: National Institutes of Health. 1990;8:1-19.

Nattinger AB, Gottlieb MS, Veum J, Yahnke D, Goodwin JS. Geographic variation in the use of breast conserving treatment for breast cancer. *New England Journal of Medicine*. 1992; 326:1102-7.

Nattinger AB, Gottlieb MS, Hoffman RG, Walker AP, Goodwin JS. Minimal increase in use of breast-conserving surgery from 1986 to 1990. *Medical Care*. 1996;34:479-489.

Newschaffer CJ, Bush TL, Penberthy LT. Comorbidity measurement in elderly female breast cancer patients with administrative and medical records data. *Journal of Clinical Epidemiology*. 1997;50:725-733.

Nyström L, Rutqvist LE, Wall S, et al. Breast cancer screening with mammography: overview of Swedish randomised trials. *The Lancet*. 1993;341:973-978.

O'Connor G, Plume S, Olmstead E, et al. A regional intervention to improve the hospital mortality associated with coronary artery bypass graft surgery. *Journal of the American Medical Association*. 1996;275:841-846.

Olson AL, Boyle WE, Evans MW, Zug LA. Overall function in rural childhood cancer survivors. The role of social competence and emotional health. *Clinical Pediatrics*. 1993;32:334-342.

Osborne CK, Clark GM, Ravdin PM. Adjuvant systemic therapy of primary breast cancer in Harris JR, Morrow M, Lippman ME, and Hellman S, eds. *Diseases of the Breast*. Lippincott- Raven Publishers, New York, 1996, pp 548-578.

Osteen RT, Cady B, Chmiel S, Clive RE, et al. 1991 national survey of carcinoma of the breast by the Commision on Cancer. 1994;178:213-219.

Osteen RT, Wincehter DP, Hussey DH, Clive RE, et al. Insurance coverage of patients with breast cancer in the 1991 Commision on Cancer patient care evaluation study. *Annals of Surgical Oncology*. 1994;1:462-467.

Osuch JR. Abnormalities on Physical Examination in . Harris JR, Morrow M, Lippman ME, and Hellman S, eds. *Diseases of the Breast*. Lippincott- Raven Publishers, New York. 1996, 110-114.

Pannek J, Rittenhouse HG, Chan DW, Epstein JI, Walsh PC, Partin AW. The use of percent free prostate specific antigen for staging clinically localized prostate cancer. *Journal of Urology*. 1998;159:1238-1242.

Pannek J, Subong EN, Jones KA, et al. The role of free/total prostate-specific antigen ratio in the prediction of final pathologic stage for men with clinically localized prostate cancer. *Urology*.1996;48(6A Suppl):51-54.

Parker SL, Tong T, Bolden S, Wingo PA. Cancer statistics. *Ca: a Cancer Journal for Clinicians*. 1997; 47: 5-27. Erratum. 1997;47:68.

Partin AW, Kattan MW, Subong EN, Walsh PC, Wojno KJ, Oesterling JE, Scardino PT, Pearson JD. Combination of prostate-specific antigen, clinical stage, and Gleason score to predict pathological stage of localized prostate cancer. A multi-institutional update. *Journal of the American Medical Association*. 1997;277:1445-1451.

Partin AW, Yoo J, Carter HB, et al. The use of prostate specific antigen, clinical stage and Gleason score to predict pathological stage in men with localized prostate cancer. *Journal of Urology*. 1993;150:110-114.

Physicians Insurers Association of America. The breast cancer study. Washington D.C., 1995.

Polednak PA. Breast cancer in black and white women in New York State. Case distribution and incidence rates by clinical stage at diagnosis. *Cancer*. 1986;58:807-815.

Potosky AL, Merrill RM, Riley GF, et al. *Journal of the National Cancer Institute.* 1997;9:1683-1691.

Potosky AL, Miller BA, Albertson PC, Kramer BS. The role of increasing detection in the rising incidence of prostate cancer. *Journal of the American Medical Association*. 1995;273:548-52.

Powell IJ, Heilbrun LK, Sakr W, Grignon D, Montie J, Novallo M, Smith D, Pontes JE. The predictive value of race as a clinical prognostic factor among patients with clinically localized prostate cancer: a multivariate analysis of positive surgical margins. *Urology*.1997;49:726-731.

Reichheld FF. Learning from customer defections: customers can help point out what is wrong with services. *Harvard Business Review*. 1996;74:56(14 pages).

Reifel J, Ganz P. Clinical applications of quality of life research. In *Cancer Policy*. Norwell, MA: Kluwer Academic Publishers, in press.

Ries LAG, Kosary CL, Hankey BF, Miller BA, Edwards BK, eds. *SEER Cancer Statistics Review, 1973-1995*. Bethesda, MD: National Cancer Institute;1998.

Riley GF, Feuer EJ, Lubitz JD. Disenrollment of Medicare cancer patients from health maintenance organizations. *Medical Care.* 1996;34:826-836.

Roberts MM, Alexander FE, Anderson TJ, et al. Edinburgh trial of screening for breast cancer: mortality at seven years. *Lancet*. 1990;335:241-246.

Rubin HR, Gandek B, Rogers WH, Kosinski M, McHorney CA, Ware JE Jr. Patients' ratings of outpatient visits in different practice settings. Results from the Medical Outcomes Study. *Journal of the American Medical Association*. 1993;270:835-840.

Samet JM, Hunt WC, Farrow DC. Determinants of receiving breast-conserving surgery. The Surveillance, Epidemiology, and End Results Program, 1983-1986. *Cancer*. 1994;73:2344-2351.

Sarrazin D, Le M, Rouesse J, et al. Conservative treatment versus mastectomy in breast cancer tumors iwht macroscopic diameter of 20 millimeters or less: the experience of the Institut Gustave-Roussy. *Cancer*. 1984;53:1209-1213.

Schatzkin A, Freedman LS, Dorgan J, McShane LM, Schiffman MH, Dawsey SM. Surrogate end points in cancer research: a critique. *Cancer Epidemiology, Biomarkers & Prevention*. 1996;5:947-953.

Schipper H, Clinch J, McMurray A, et al. Measuring the quality of life of cancer patients: the Functional Living Index-Cancer: development and validation. *Journal of Clinical Oncology*. 1984;2:472-483.

Schuster MA, Asch SM, McGlynn EA, Kerr EA, Hardy AM, Gifford DS. Development of a quality of care measurement system for children and adolescents. *Archives of Pediatrics and Adolescent Medicine*. 1997;151:1085-1092.

Schuster MA, McGlynn EA, Brook RH. "How Good Is the Quality of Health Care in the United States?" *Milbank Quarterly*. 1998;76:in press.

SEER Website, http://www-seer.ims.nci.nih.gov, accessed 6/29/98.

Segal M. Mammography facilities must meet quality standards. *FDA Consumer*. March 1994; p 8-12.

Shapiro S, Venet W, Strax P, et al. Periodic screening for breast cancer -- The Health Insurance Plan project and its sequelae, 1963-1986. Baltimore, Johns Hopkins University Press, 1988.

Starfield B, Riley AW, Green BF, et al. The Adolescent Child Health and Illness Profile. A Population-Based Measure of Health. *Medical Care.* 1995;33:553.

Stein RE, Jessop DJ. Functional status II (R). A measure of child health status. *Medical Care.* 1990;28:1041. Stein RE, Jessop DJ. Erratum. *Medical Care*. 1991;29:490.

Stiller CA. Centralisation of treatment and survival rates for cancer. *Archives of Disease in Childhood*. 1988;63:23-30.

Stolier AJ. Stereotactic breast biopsy: a surgical series. *Journal of the American College of Surgeons.* 1997;185:224-228.

Strayer F, Kisker T, Fethke C. Cost-effectiveness of a shared-management delivery system for the care of children with cancer. *Pediatrics*. 1980;66:907-911.

Tabar L, Fagerberg G, Chen H-H, et al. Efficacy of breast cancer screening by age. New results from the Swedish two-county trial. *Cancer*. 1995;75:2507-2517.

Talamonti MS, Morrow M. The abnormal mammogram in Harris JR, Morrow M, Lippman ME, and Hellman S, eds. *Diseases of the Breast*.. Lippincott- Raven Publishers, New York, 1996, pp 114-122.

U.S. Preventive Services Task Force. *Guide to clinical preventive services*. Alexandria, VA: International Medical Publishing. 1996.

Veronesi U, Saccozzi R, Del Vecchio M, et al. Comparing radical mastectomy with quandrantectomy, axillary dissection, and radiotherapy in patients with small cancers of the breast. *New England Journal of Medicine*. 1981;305:6-11.

von Eschenbach A, Ho R, Murphy GP, Cunningham M, Lins N. American Cancer Society guideline for the early detection of prostate cancer: update 1997. *CA: Cancer Journal for Clinicians*. 1997;47:261-264.

Wasson JH, Cushman CC, Bruskewitz RC, Littenberg B, Mulley AG Jr, Wennberg JE. A structured literature review of treatment for localized prostate cancer. Prostate Disease Patient Outcome Research. *Archives of Family Medicine.* 1993;2:487-93.

Weeks JC. Outcomes assessment in the NCCN. *Oncology*. 1997;11(11A): 137-140.

Wells WAS, Carney PA, Eliassen MS, Tosteson AN, Greenberg ER. Statewide study of diagnostic agreement in breast pathology. *Journal of the National Cancer Institute.* 1998; 90:142-145.

Wennberg JE and McAndrew Cooper M. *The Dartmouth Atlas of Health Care in the United States*. Chicago, Illinois: American Hospital Publishing, Inc. 1996. pp. 128-129.

Wilson NJ, Kizer KW. Oncology management by the "new" Veterans Health Administration. *Cancer*, 1998;82(10; Supple): 2003-2009.

Wincester DP, Osteen RT, Menck HR. The National Cancer Data Base report on breast carcinoma characteristics and outcome in relation to age. *Cancer*. 1996;78:1838-1843.

Wingo PA, Landis S, Ries LAG. An adjustment to the 1997 estimate for new prostate cancer cases. *Cancer*. 1997;80:1810-3.

Young PC, Wasserman RC, McAullife T, Long J, Hagan JF, Heath B. Why families change pediatricians. Factors causing dissatisfaction with pediatric care. *American Journal of Diseases of Children*. 1985;139:683-686.

Young WW, Marks SM, Kohler SA, and Hsu AY. Dissemination of clinical results. Mastectomy versus lumpectomy and radiation therapy. *Medical Care*. 1996. 34:1003-1017.

Zaloznik AJ. Breast cancer stage at diagnosis: Caucasians versus Afro-Americans. *Breast Cancer Research and Treatment*. 1995;34:195-198.

Zaloznik AJ. Breast cancer stage at diagnosis: Caucasians versus Hispanics. *Breast Cancer Reseatch and Treatment.* 1997;42:121-124.

Zippin C, Lum D, Hankey BF. Completeness of hospital cancer case reporting from the SEER program of the National Cancer Institute. *Cancer*. 1995;76: 2343-2350.

Acknowledgments

We are indebted to Robert H. Brook, M.D., Sc.D., and Jay G. Tassin, M.D., for comments on a draft of this paper, and to Yuko Sano, A.B., and Cung B. Pham, B.A., for research assistance. This work was supported in part by the National Cancer Policy Board, Institute of Medicine, National Academy of Sciences, and the Bing Fund.