

INTRODUCTION

Developing and implementing a valid system of quality assessment is essential for effective functioning of the health care system. Although a number of groups have produced quality assessment tools, these tools typically suffer from a variety of limitations. Information is obtained on only a few dimensions of quality, the tools rely exclusively on administrative data, they examine quality only for users of services rather than the population, or they fail to provide a scientific basis for the quality indicators.

Under funding from public and private sponsors, including the Health Care Financing Administration (HCFA), the Agency for Healthcare Research and Quality (AHRQ), the California HealthCare Foundation, and the Robert Wood Johnson Foundation (RWJ), RAND has developed and tested a comprehensive, clinically based system for assessing quality of care for children and adults. We call this system QA Tools.

In this introduction, we discuss how the clinical areas were selected, how the indicators were chosen, and what is included in the overall system. We then describe in detail how we developed the indicators for children and adolescents.

ADVANTAGES OF THE QA TOOLS SYSTEM

QA Tools is a comprehensive, clinically based system for assessing the quality of care for children and adults. The indicators cover 46 clinical areas and all four functions of medicine including screening, diagnosis, treatment, and follow-up. The indicators also cover a variety of modes of providing care, such as history, physical examination, laboratory study, medication, and other interventions and contacts. Initial development of indicators for each clinical area was based on a review of the literature.

The QA Tools system addresses many limitations of current quality assessment tools by offering the following:

- They are clinically detailed and require data typically found in medical records rather than just relying exclusively on data from administrative records.

- They examine quality for a population-based sample rather than for a more restricted sample of those who use care or have insurance.
- They document the scientific basis for developing and choosing the indicators.
- The QA Tools system is designed to target populations vulnerable to underutilization.
- Because of the comprehensiveness of the system, it is difficult for health care organizations to focus on a few indicators to increase their quality scores.
- QA Tools is a system that can be effective for both internal and external quality reviews. Health care organizations can use the system in order to improve the overall quality of the care provided.
- Because of the simple summary scores that will be produced, it will be an important tool for purchasers and consumers who are making choices about health care coverage and which provider to see.

Given its comprehensiveness, the QA Tools system contrasts with *leading indicators*, the most common approach to quality measurement in use today. Under the leading indicators approach, three to five specific quality measures are selected across a few domains (for example, rates of mammography screening, prevalence of the use of beta blockers among persons who have had a heart attack, and appropriateness of hysterectomy).

Leading indicators may work well for drawing general conclusions about quality when they correlate highly with other similar but unmeasured interventions and when repeated measurement and public reporting does not change the relationship of those indicators to the related interventions. However, to date no real evaluation of the utility of leading indicators in assessing health system performance has been done. We also do not know whether the selected indicators currently in use consistently represent other unmeasured practices.

By contrast, a comprehensive system can represent different dimensions of quality of care delivery by using a large number of measures applied to a population of interest and aggregated to produce index scores to draw conclusions about quality. A comprehensive system works well when evidence exists of variability within and between the diagnosis and management of different conditions and when the question being asked is framed at a high

level (for instance, how well is the health system helping the population stay healthy, or how much of a problem does underuse present?).

In the 46 clinical areas they encompass, the QA Tools adequately represent scientific and expert judgment on what constitutes quality care. However, both the science and the practice of medicine continue to evolve. For the QA Tools to remain a valid tool for quality assessment over time, the scientific evidence in each area needs to be reviewed annually to determine if new evidence warrants modifying the indicators and/or clinical areas included in the system.

SELECTING CLINICAL AREAS FOR THE QA TOOLS

We reviewed Vital Statistics, the National Health Interview Survey, the National Hospital Discharge Survey, and the National Ambulatory Medical Care Survey to identify the leading causes of morbidity and mortality and the most common reasons for physician visits in the United States. We examined statistics for different age and gender groups in the population (0-1, 1-5, 6-11, 12-17, 18-50 [men and women], 50-64, 65-75, over 75).

We selected topics that reflected these different areas of importance (death, disability, utilization of services) and that covered preventive care as well as care for acute and chronic conditions. In addition, we consulted with a variety of experts to identify areas that are important to these various populations but that may be underrepresented in national data sets (for example, mental health problems). Finally, we sought to select enough clinical areas to represent a majority of the health care delivery system.

Table I.1 lists the 46 clinical areas included in the QA Tools system by population group; 20 include indicators for children and 36 for adults. The clinical areas, broadly defined, represent about 55 percent of the reasons for ambulatory care visits among children, 50 percent of the reasons for ambulatory care visits for the entire population, and 46 percent of the reasons for hospitalization among adults.

Note: Table I.1 reflects the clinical areas that were included in the system currently being tested. Several clinical areas (e.g., lung cancer, sickle cell disease) for which indicators were developed were not incorporated into the current tool due to budgetary constraints.

Table I.1

Clinical Areas in QA Tools System By Covered Population Group

Clinical Areas	Children	Adults
Acne	X	
Adolescent preventive services	X	
Adult screening and prevention		X
Alcohol dependence		X
Allergic rhinitis	X	
Asthma	X	X
Atrial fibrillation		X
Attention deficit/hyperactivity disorder	X	
Benign prostatic hyperplasia		X
Breast cancer		X
Cataracts		X
Cerebrovascular disease		X
Cervical cancer		X
Cesarean delivery	X	X
Chronic obstructive pulmonary disease		X
Colorectal cancer		X
Congestive heart failure		X
Coronary artery disease		X
Depression	X	X
Developmental screening	X	
Diabetes Mellitus	X	X
Diarrheal disease	X	
Family planning and contraception	X	X
Fever of unknown origin	X	
Headache		X
Hip fracture		X
Hormone replacement therapy		X
Human immunodeficiency virus		X
Hyperlipidemia		X
Hypertension		X
Immunizations	X	X
Low back pain		X
Orthopedic conditions		X
Osteoarthritis		X
Otitis media	X	
Pain management for cancer		X
Peptic ulcer disease & dyspepsia		X
Pneumonia		X
Prenatal care and delivery	X	X
Prostate cancer		X
Tuberculosis	X	X
Upper respiratory tract infections	X	
Urinary tract infections	X	X
Uterine bleeding and hysterectomy		X
Vaginitis and sexually transmitted diseases	X	X
Well child care	X	
Total number of clinical areas	20	36

SELECTING QUALITY INDICATORS

In this section, we describe the process by which indicators were chosen for inclusion in the QA Tools system. This process involved RAND staff drafting proposed indicators based on a review of the pertinent clinical literature and expert panel review of those indicators.

Literature Review

For each clinical area chosen, we reviewed the scientific literature for evidence that effective methods of prevention, screening, diagnosis, treatment, and follow-up existed (Kerr et al., 2000a; Kerr et al., 2000b; McGlynn et al., 2000a; McGlynn et al., 2000b). We explicitly examined the continuum of care in each clinical area. RAND staff drafted indicators that

- addressed an intervention with potential health benefits for the patient
- were supported by scientific evidence or formal professional consensus (guidelines, for example)
- can be significantly influenced by the health care delivery system
- can be assessed from available sources of information, primarily the medical record.

The literature review process varied slightly for each clinical area, but the basic strategy involved the following:

- Identify general areas in which quality indicators are likely to be developed.
- Review relevant textbooks and review articles.
- Conduct a targeted MEDLINE search on specific topics related to the probable indicator areas.

The levels of evidence for each indicator were assigned to three categories: randomized clinical trial; nonrandomized controlled trials, cohort or case analysis, or multiple time series; and textbooks, opinions, or descriptive studies. For each proposed indicator, staff noted the highest level of evidence supporting the indicator.

Because of the breadth of topics for which we were developing indicators, some of the literature reviews relied exclusively on textbooks and review articles. Nonetheless, we believe that the reviews adequately summarize clinical opinion and key research at the time that they were conducted. The

literature reviews used to develop quality indicators for children and adolescents, and for women, were conducted between January and July 1995. The reviews for general medical conditions, oncologic conditions, and cardiopulmonary conditions were conducted between November 1996 and July 1997.

For each clinical area, we wrote a summary of the scientific evidence and developed tables of the proposed indicators that included the level of evidence, specific studies in support of the indicator, and the clinical rationale for the indicator. Because the organization of care delivery is changing so rapidly, we drafted indicators that were not in most cases inextricably linked to the place where the care was provided.

Types of Indicators

Quality of care is usually determined with three types of measures:

- **Structural measures** include characteristics of clinicians (for instance, board certification or years of experience), organizations (for instance, staffing patterns or types of equipment available), and patients (for instance, type of insurance or severity of illness).
- **Process measures** include the ways in which clinicians and patients interact and the appropriateness of medical treatment for a specific patient.
- **Outcomes measures** include changes in patients' current and future health status, including health-related quality of life and satisfaction with care.

The indicators included in the QA Tools system are primarily process indicators. We deliberately chose such indicators because the system was designed to assess care for which we can hold providers responsible. However, we collect data on a number of intermediate outcomes measures (for example, glycosylated hemoglobin, blood pressure, and cholesterol) that could be used to construct intermediate clinical outcomes indicators.

In many instances, the measures included in the QA Tools system are used to determine whether interventions have been provided in response to poor performance on such measures (for instance, whether persons who fail to control their blood sugar on dietary therapy are offered oral hypoglycemic therapy).

The Expert Panel Process

We convened expert panels to evaluate the indicators and to make final selections using the RAND/UCLA Appropriateness Method, a modified Delphi method developed at RAND and UCLA (Brook 1994). In general, the method quantitatively assesses the expert judgment of a group of clinicians regarding the indicators by using a scale with values ranging from 1 to 9.

The method is iterative with two rounds of anonymous ratings of the indicators by the panel and a face-to-face group discussion between rounds. Each panelist has equal weight in determining the final result: the quality indicators that will be included in the QA Tools system.

The RAND/UCLA Appropriateness Method has been shown to have a reproducibility consistent with that of well accepted diagnostic tests such as the interpretation of coronary angiography and screening mammography (Shekelle et al., 1998a). It has also been shown to have content, construct, and predictive validity in other applications (Brook, 1994; Shekelle et al., 1998b; Kravitz et al., 1995; Selby et al., 1996).

Approximately six weeks before the panel meeting, we sent panelists the reviews of the literature, the staff-proposed quality indicators, and separate rating sheets for each clinical area. We asked the panelists to examine the literature review and rate each indicator on a nine-point scale on each of two dimensions: validity and feasibility.

A quality indicator is defined as valid if:

1. Adequate scientific evidence or professional consensus exists supporting the indicator.
2. There are identifiable health benefits to patients who receive care specified by the indicator.
3. Based on the panelists' professional experience, health professionals with significantly higher rates of adherence to an indicator would be considered higher quality providers
4. The majority of factors that determine adherence to an indicator are under the control of the health professional (or are subject to influence by the health professional—for example, smoking cessation).

Ratings of 1-3 mean that the indicator is not a valid criterion for evaluating quality. Ratings of 4-6 mean that the indicator is an uncertain or

equivocal criterion for evaluating quality. Ratings of 7-9 mean that the indicator is clearly a valid criterion for evaluating quality.

A quality indicator is defined as feasible if:

1. The information necessary to determine adherence is likely to be found in a typical medical record.
2. Estimates of adherence to the indicator based on medical record data are likely to be reliable and unbiased.
3. Failure to document relevant information about the indicator is itself a marker for poor quality.

Ratings of 1-3 mean that it is not feasible to use the indicator for evaluating quality. Ratings of 4-6 mean that there will be considerable variability in the feasibility of using the indicator to evaluate quality. Ratings of 7-9 mean that it is clearly feasible to use the indicator for evaluating quality.

The first round of indicators was rated by the panelists individually in their own offices. The indicators were returned to RAND staff and the results of the first round were summarized. We encouraged panelists to comment on the literature reviews, the definitions of key terms, and the indicators. We also encouraged them to suggest additions or deletions to the indicators.

At the panel meeting, participants discussed each clinical area in turn, focusing on the evidence, or lack thereof, that supports or refutes each indicator and the panelists' prior validity rankings. Panelists had before them the summary of the panel's first round ratings and a confidential reminder of their own ratings.

The summary consisted of a printout of the rating sheet with the distribution of ratings by panelists displayed above the rating line (without revealing the identity of the panelists) and a caret (^) marking the individual panelist's own rating in the first round displayed below the line. An example of the printout received by panelists is shown in Figure I.1.

Chapter 1 ASTHMA	Validity									Feasibility																		
DIAGNOSIS																												
3. Spirometry should be measured in patients with chronic asthma at least every 2 years.	1	1	2	3	1	1				1	2	3	4	5	6	7	8	9	3	4	2							(1- 2)
									^																		^	
TREATMENT																												
7. Patients requiring chronic treatment with systemic corticosteroids during any 12 month period should have been prescribed inhaled corticosteroids during the same 12 month period.										1	6	2																(3- 4)
10. All patients seen for an acute asthma exacerbation should be evaluated with a complete history including all of the following:																												
a. time of onset										2	2	2	3						2	2	1	1	3					(5- 6)
b. all current medications																												(7- 8)
c. prior hospitalizations and emergency department visits for asthma																												(9-10)
d. prior episodes of respiratory insufficiency due to asthma																												(11-12)

Scales: 1 = low validity or feasibility; 9 = high validity or feasibility

Figure I.1 - Sample Panelist Summary Rating Sheet

Panelists were encouraged to bring to the discussion any relevant published information that the literature reviews had omitted. In a few cases, they supplied this information which was, in turn, discussed. In several cases, the indicators were reworded or otherwise clarified to better fit clinical judgment.

After further discussion, all indicators in each clinical area were re-ranked for validity. These final round rankings were analyzed in a manner similar to past applications of the RAND/UCLA Appropriateness Method (Park et al., 1986; Brook, 1994). The median panel rating and measure of dispersion were used to categorize indicators on validity.

We regarded panel members as being in *disagreement* when at least three members of the panel judged an indicator as being in the highest tertile of validity (that is, having a rating of 7, 8, or 9) and three members rated it as being in the lowest tertile of validity (1, 2, or 3) (Brook, 1994). Indicators with a median validity rating of 7 or higher without disagreement were included in the system.

We also obtained ratings from the panelists about the feasibility of obtaining the data necessary to score the indicators from medical. This was done to make explicit that failure to document key variables required to score an indicator would be treated as though the recommended care was not provided.

Although we do not intend for quality assessment to impose significant additional documentation burdens, we wanted the panel to acknowledge that documentation itself is an element of quality particularly when patients are treated by a team of health professionals. Because of the variability in documentation patterns and the opportunity to empirically evaluate feasibility, indicators with a median feasibility rating of 4 and higher were accepted into the system. Indicators had to satisfy both the validity and feasibility criteria.

Five expert panels were convened on the topics of children's care, care for women 18-50, general medicine for adults, oncologic conditions and HIV, and cardiopulmonary conditions.

The dates on which the panels were conducted are shown in Table I.2.

Table I.2
Dates Expert Panels Convened

Children	October 1995
Women	November 1995
Cardiopulmonary	September 1997
Oncology/HIV	October 1997
General Medicine	November 1997

Tables I.3 through I.6 summarize the distribution of indicators by level of evidence, type of care (preventive, acute, chronic), function of medicine (screening, diagnosis, treatment, follow-up, continuity), and modality (for

example, history, physical examination, laboratory test, medication) (Malin et al., 2000; Schuster et al., 1997).

The categories were selected by the research team and reflect terminology commonly used by health services researchers to describe different aspects of health service delivery. The categories also reflect the areas in which we intend to develop aggregate quality of care scores. However, a significant benefit of the QA Tools system is its adaptability to other frameworks.

Note: In the following tables, the figures in some columns may not total exactly 100 percent due to the rounding of fractional numbers.

Table I.3
Distribution of Indicators (%) by Level of Evidence

Level of Evidence	Children	Women	Cancer/HIV	Cardio- pulmonary	General Medicine
Randomized trials	11	22	22	18	23
Nonrandomized trials	6	16	37	4	17
Descriptive studies	72	59	26	71	57
Added by panel	12	4	15	7	4
Total	101	101	100	100	101

Table I.4
Distribution of Indicators (%) by Type of Care

Type of Care	Children	Women	Cancer/HIV	Cardio-pulmonary	General Medicine
Preventive	30	11	20	3	18
Acute	36	49	7	26	38
Chronic	34	41	74	71	44
Total	100	101	101	100	100

Table I.5
Distribution of Indicators (%) by Function of Medicine

Function of Medicine	Children	Women	Cancer/HIV	Cardio-pulmonary	General Medicine
Screening	23	18	9	3	12
Diagnosis	31	30	27	54	41
Treatment	36	43	53	36	41
Follow-up	10	12	10	8	6
Total	100	103	99	101	100

Table I.6
Distribution of Indicators (%) by Modality

Modality	Children	Women	Cancer/HIV	Cardio- pulmonary	General Medicine
History	19	18	4	11	23
Physical	19	10	5	21	15
Lab/Radiology	21	23	24	23	18
Medication	25	29	25	25	26
Other	17	19	42	20	17
Total	101	99	100	100	99

DEVELOPING QUALITY INDICATORS FOR ONCOLOGIC CONDITIONS AND HIV

We now describe in more detail the process by which we developed quality indicators for oncologic conditions and HIV.

Selecting Clinical Areas

We began our selection of clinical areas by examining national data sources to identify the leading causes of mortality, morbidity, and functional limitation among adult men and women. The principal data sources for this review were Vital Statistics, the National Health Interview Survey (NHIS), the National Ambulatory Medical Care Survey (NAMCS), and the National Hospital Discharge Survey (NHDS).

From these data sources, we selected the conditions that represent the top causes of mortality, hospitalization, and outpatient visits. This process led to the selection of some areas that were developed for the women's care panel (McGlynn et al., 2000b).

To facilitate the review and rating process, we grouped the selected areas into three categories: cardiopulmonary conditions, oncologic conditions and HIV, and general medical conditions. Table I.7 lists the clinical areas covered by each of these categories. "Cancer Pain and Palliation" was not

among the originally selected clinical areas, but was added during the panel process as a result of strong recommendations from several oncology panelists.

Table I.7
Clinical Areas Covered by Each Expert Panel

Cardiopulmonary (N=12)	Oncology and HIV (N=11)	General Medicine (N=22)
Asthma*	Breast Cancer Screening	Acne*
Atrial Fibrillation	Breast Cancer Diagnosis and Treatment*	Alcohol Dependence*
Cerebrovascular Disease	Cervical Cancer Screening*	Allergic Rhinitis*
Chronic Obstructive Pulmonary Disease	Colorectal Cancer Screening	Benign Prostatic Hyperplasia
Cigarette Counseling*	Colorectal Cancer Diagnosis and Treatment	Cataracts
Congestive Heart Failure	HIV Disease	Cholelithiasis
Coronary Artery Disease Diagnosis and Screening	Lung Cancer	Dementia
Coronary Artery Disease Prevention and Treatment	Prostate Cancer Screening	Depression*
Hyperlipidemia	Prostate Cancer Diagnosis and Treatment	Diabetes Mellitus*
Hypertension*	Skin Cancer Screening	Dyspepsia and Peptic Ulcer Disease
Pneumonia	Cancer Pain and Palliation	Hormone Replacement Therapy
Upper Respiratory Infections*		Headache*
		Hip Fracture
		Hysterectomy
		Inguinal Hernia
		Low Back Pain (Acute)*
		Orthopedic Conditions
		Osteoarthritis
		Preventive Care*
		Urinary Tract Infections*
		Vaginitis and Sexually Transmitted Diseases*
		Vertigo and Dizziness

* Previously addressed by the panel on quality of care for women (McGlynn et al., 2000b).

Conducting Literature Reviews

The literature reviews were conducted as described earlier in this Introduction by a team of 14 physician investigators, many of whom have clinical expertise in the conditions selected for this project. Each investigator drafted a review of the literature for his or her topic area, focusing on important areas for quality measurement (as opposed to a clinical review of the literature, which would focus on clinical management) and drafted potential indicators.

Every indicator table was then reviewed by Drs. Asch or Kerr for content, consistency, and the likely availability of information necessary to score adherence to the indicator from the medical record. On a few occasions, when questions remained even after detailed literature review, we requested that a clinical leader in the field read and comment on the draft review and indicators.

In addition, the physician investigators updated the 16 clinical areas carried over from the women's care panel. This included reading the reviews and indicators from the women's care panel, updating the supporting literature to 1997, and modifying the pre-existing indicators as appropriate. In most cases few changes were made, but indicators were deleted if the evidence changed or if our implementation experience proved that it was not feasible to collect the data necessary to determine eligibility and/or a scoring indicator. Indicators were added if strong evidence since 1995 supported the need for a new criterion. In the clinical areas previously addressed, the expert panels for this project rated only those indicators that had been added or significantly revised (indicated by bold type in the indicator tables in the chapters that follow).

This quality assessment system was designed to encompass a substantial portion of the inpatient and ambulatory care received by the population. In order to estimate the percentage of ambulatory care visits covered by this system, we aggregated applicable ICD-9 codes into the clinical areas for which we are developing quality indicators. We then calculated the number of adult visits for each condition in the 1993 National Ambulatory Medical Care Survey (NAMCS). We used the same method to estimate the percentage of inpatient admissions accounted for by each clinical area in the 1992 National Hospital Discharge Survey.

Aggregating ICD-9 codes into the clinical areas covered by this system was an imprecise task, requiring a rather broad definition of what is "included" in each clinical area. The 45 clinical conditions covered by this quality measurement system encompass 50 percent of all ambulatory care visits and 46 percent of non-federal inpatient hospital admissions.

Developing Indicators

In each clinical area, we developed indicators defining the explicit criteria by which quality of care would be evaluated. These indicators focus on technical processes of care for the various conditions and are organized by function: screening, diagnosis, treatment and follow-up. Although we have developed indicators across the continuum of management for each condition, we have not attempted to cover every important area or every possible clinical circumstance. The indicators were designed to apply to the average patient with the specified condition who is seeing the average physician.

Our approach makes a strong distinction between indicators of quality of care and practice guidelines (see Table I.8). While guidelines are intended to be comprehensive in scope, indicators are meant to apply to specific clinical circumstances in which there is believed to be a strong link between a measurable health care process and patient outcomes.

Indicators are not intended to measure all possible care for a condition. Furthermore, guidelines are intended to be applied prospectively at the individual patient level, whereas indicators are applied retrospectively and scored at an aggregate level. Finally, indicators must be written precisely in order to be *operationalized* (that is, to form useful measures of quality based on medical records or administrative data).

Table I.8
Clinical Guidelines versus Quality Indicators

Guidelines	Indicators
<p>Comprehensive: Cover virtually all aspects of care for a condition.</p> <p>Prescriptive: Intended to influence provider behavior prospectively at the individual patient level.</p> <p>Flexible: Intentionally allow room for clinical judgment and interpretation.</p>	<p>Targeted: Apply to specific clinical circumstances in which there is evidence of a process-outcome link.</p> <p>Observational: Measure past provider behavior at an aggregate level.</p> <p>Operational: Precise language that can be applied systematically to medical records or administrative data.</p>

The indicator tables at the end of each chapter of this book

- note the population to whom the indicators apply
- list the indicators themselves
- provide a "grade" for the strength of the evidence that supports each indicator
- list the specific literature used to support each indicator
- provide a statement of the health benefits of complying with each indicator
- include comments to further explain the purpose or reasoning behind each indicator.

Selecting Panel Participants

We requested nominations for potential expert panel participants from the relevant specialty societies for oncology and HIV: the American College of Physicians, American Academy of Family Physicians, American Geriatrics Society, American Cancer Society, American Society of Clinical Oncology, Infectious Disease Society of America, and the Society of General Internal Medicine. We received a total of 206 nominations for the panels on general internal medicine, oncology and HIV, and cardiopulmonary conditions.

Each nominee was sent a letter summarizing the purpose of the project and indicating which group recommended them. Interested candidates were asked to return a curriculum vitae and calendar with available dates. We received positive responses from 156 potential panelists. The quality of the recommended panelists was excellent.

We sought to ensure that each panel was diverse with respect to type of practice (academic, private practice, managed care organizational practice), geographic location, gender, and specialty. The oncology panel included seven oncologists and two general internists. Dr. Lodovico Balducci, an oncologist, was selected by RAND staff to chair this panel (see the Acknowledgements earlier in this book for the list of panelists).

Selecting the Final Indicators

The panel process was conducted as described earlier in this Introduction.

During the course of the Oncology and HIV panel meeting, several panelists noted the rapidity with which clinical practice in this field is

changing. Therefore, process of care criteria that represented standard practice at the time of the panel meeting may soon become obsolete as new practices are found to be efficacious.

For many patients, the best option may be enrollment in a clinical trial of a new therapy (where standard therapy is the randomized alternative). To ensure that the importance of clinical trials is recognized, we have included enrollment in such trials, with documentation of informed consent, as meeting the intent of several oncology indicators for which panelists felt this option was appropriate. For all other indicators included in the system, patients participating in a relevant clinical trial will be considered ineligible for the intervention specified in the indicator(s) and will not be included in the scoring.

Analyzing the Final Set of Indicators

A total of 145 quality indicators (including subparts) were reviewed by the oncology and HIV expert panel. All 11 indicators retained from the women's care panel were accepted by the oncology/HIV panel on the basis of the women's panel's ratings. Six indicators were deleted before the final panel ratings in response to revisions suggested by panelists prior to or during the panel meeting. Of the remaining 128 indicators that received final ratings from this panel, 20 were added by the oncology and HIV care panel itself. This occurred either when panelists agreed that a new indicator should be written to cover an important topic, or, more frequently, as a result of splitting a staff-proposed indicator.

The panel accepted 106 (83%) of the 128 indicators it rated. Twenty-two indicators (17%) were dropped due to low ratings: 17 for low validity scores, 3 for substantial disagreement on validity, and 2 for low feasibility scores.

Table I.9 summarizes the disposition of all 145 proposed oncology and HIV quality indicators by the strength of their supporting evidence. The final set consists of 117 indicators (106 rated by this panel and 11 approved based on ratings by the women's panel), or 81 percent of those proposed. Table I.9 reveals that indicators that are not based on randomized clinical trials (that is, Level II and Level III indicators) were much more likely to be rejected by the panel. Similarly, indicators proposed by the panelists themselves fared

poorly relative to those with Level I evidence. This pattern has been observed consistently across several RAND quality of care panels.

Table I.9
Disposition of Proposed Oncology and HIV Quality Indicators
by Strength of Evidence

Strength of Evidence	Total Proposed	Indicator Disposition			
		Accepted	Retained from Women's Panel	Drop before Rating	Drop Due to Low Rating
I. Randomized controlled trials	27 (100%)	26 (96%)	0 (0%)	0 (0%)	1 (4%)
II. Non-randomized trials	57 (100%)	41 (72%)	2 (4%)	2 (4%)	12 (21%)
III. Opinions, descriptive studies, or textbooks	41 (100%)	22 (54%)	9 (22%)	4 (10%)	6 (15%)
IV. Added by Clinical Panel	20 (100%)	17 (85%)	0 (0%)	0 (0%)	3 (15%)
Total	145 (100%)	106 (73%)	11 (8%)	6 (4%)	22 (15%)

The summary ratings sheets for oncology and HIV are shown in Appendix A of this book.

Figure I.2 provides an example of a final summary rating sheet. The chapter number and clinical condition are shown in the top left margin. The rating bar is numbered from 1 to 9, indicating the range of possible responses. The number shown above each of the responses in the rating bar indicates how many panelists provided that particular rating for the indicator. Below the score distribution, in parentheses, the median and the absolute deviation from the media are listed. Each dimension is assigned an A for "Agreement", D for "Disagreement", or I for "Indeterminate" based on the score distribution.

Note: We recommend caution when reviewing the ratings for each indicator. The overall median does not tell us anything about the extent to which the

indicators occur in clinical practice. To determine that, actual clinical data to assess the indicators must be collected and analyzed.

Chapter 6 HIV DISEASE		Validity	Feasibility
SCREENING AND PREVENTION			
1. HIV+ patients should be offered PCP prophylaxis within one month of meeting any of the following conditions:			
a. CD4 count dropping below 200	1 2 3 4 5 6 7 8 9	1 8 (9.0, 0.1, A)	3 6 (9.0, 0.3, A)
b. Thrush	2 1 2 1 1 2	2	1 1 1 2 2
c. Completion of active treatment of PCP	1 2 3 4 5 6 7 8 9	2 7 (9.0, 0.2, A)	3 6 (9.0, 0.3, A)
d. CD4 below 15%	1 2 3 4 5 6 7 8 9	1 1 1 6 (9.0, 0.7, A)	3 6 (9.0, 0.3, A)
2. HIV+ patients who do not have active TB and who have not ever previously received TB prophylaxis should be offered TB prophylaxis within one month of meeting any of following conditions:			
a. Current PPD > 5 mm	1 2 3 4 5 6 7 8 9	1 3 5 (9.0, 0.6, A)	1 3 5 (9.0, 0.6, A)
b. Provider noting that patient has had PPD > 5 mm administered at anytime since HIV diagnosis	1 2 3 4 5 6 7 8 9	2 4 3 (8.0, 0.6, A)	1 1 3 1 3 (7.0, 1.2, A)
c. Contact with person with active TB	2 1 1 1 2 2	2	1 2 1 3
3. HIV+ patients who do not have active toxoplasmosis should be offered toxoplasmosis prophylaxis within one month of meeting either of the following conditions:	1 2 3 4 5 6 7 8 9	(7.0, 2.7, D)	(4.0, 1.6, I)
- Toxo IgG positive and CD4 count dropping below 100			
- Completion of therapy for active toxoplasmosis	1 2 3 4 5 6 7 8 9	3 6 (9.0, 0.7, A)	2 2 5 (9.0, 0.7, A)
4. HIV+ patients should have toxoplasmosis serology documented.	1 2 3 4 5 6 7 8 9	4 1 4 (8.0, 0.9, A)	3 4 2 (8.0, 0.6, A)
5. HIV+ patients should be offered MAC prophylaxis within one month of a CD4 count dropping below 50.	1 2 3 4 5 6 7 8 9	2 2 5 (9.0, 0.7, A)	3 2 4 (8.0, 0.8, A)
6. HIV+ patients with a lowest recorded CD4 > 200 should have a documented pneumovax.	1 2 3 4 5 6 7 8 9	1 3 3 2 (8.0, 0.8, A)	1 1 5 2 (7.0, 1.1, A)
7. HIV+ patients with a lowest recorded CD4 count of less than 100 should have had a yearly dilated fundoscopic exam.	1 2 3 4 5 6 7 8 9	4 2 3 (8.0, 0.8, A)	1 2 3 3 (8.0, 0.8, A)
8. HIV+ patients should have a VDRL or RPR documented in the chart.	1 2 3 4 5 6 7 8 9	3 2 4 (8.0, 0.8, A)	2 3 4 (8.0, 0.7, A)
9. Sexually active HIV+ patients should be offered a VDRL/RPR annually.	7 1 1	5 1 1 2 (1.0, 1.0, A)	1 2 3 4 5 6 7 8 9 (1.0, 1.2, A)

Scales: 1 = low validity or feasibility; 9 = high validity or feasibility

Figure I.2 - Sample Rating Results Sheet

The tables in Appendix B show the changes made to each indicator during the panel process, the reasons for those changes, and the final disposition of each indicator. Wherever possible, we have tried to briefly summarize the discussion that led the panel to either modify or drop indicators. These explanations are based on extensive notes taken by RAND staff during the panel process, but should not be considered representative of the views of all of the panelists, nor of any individual.

Because the final quality assessment system will produce aggregate scores for various dimensions of health care, it is useful to examine the distribution of the final indicators across some of these dimensions. Table I.10 summarizes the distribution of quality indicators by type of care (preventive, acute, and chronic), the function of the medical care provided (screening, diagnosis, treatment, and follow-up), and the modality by which care is delivered (history, physical examination, laboratory or radiologic study, medication, other interventions,¹ and other contacts²). Indicators were assigned to only one type of care, but could have up to two functions and three modalities.

Indicators with more than one function or modality were allocated fractionally across categories. For example, one indicator states, "For patients who present with a complaint of sore throat, a history/physical exam should document presence or absence of: a) fever; b) tonsillar exudate; c) anterior cervical adenopathy." This indicator was allocated 50 percent to the history modality and 50 percent to the physical examination modality.

¹ Other interventions include counseling, education, procedures, and surgery.

² Other contacts include general follow-up visit or phone call, referral to subspecialist, or hospitalization.

Table I.10

Distribution of Final Oncology and HIV Quality Indicators by Type of Care, Function, and Modality

	Number of Indicators	Percent of Indicators
Type		
Preventive	23	20%
Acute	8	7%
Chronic	86	74%
Function		
Screening	11	9%
Diagnosis	32	27%
Treatment	62	53%
Follow-up	12	10%
Modality*		
History	5	4%
Physical Examination	6	5%
Laboratory or Radiologic Study	28	24%
Medication	29	25%
Other Intervention	50	43%
Other Contact	0	0%
Total	117	100%

* Total does not sum to 117 due to rounding.

CONCLUSION

This report provides the foundation for a broad set of quality indicators covering oncology and HIV health care. The final indicators presented here cover a variety of clinical conditions, span a range of clinical functions and modalities, and are rated by the level of evidence in the supporting literature. When combined with the indicators approved by the women's care, child and adolescent care, cardiopulmonary, and general medicine expert panels, the complete quality assessment system will be more comprehensive than any quality assessment system in use today.

The comprehensive nature of this system is demonstrated by the broad scope of the indicators. Of the 145 indicators reviewed by the oncology and HIV expert panel, 117 (81%) were retained. These indicators cover a mix of preventive, acute, and chronic care. However, given the chronic nature of the

conditions covered by this panel, a large proportion of the indicators (74%) fall into the chronic care category. They address all four functions of medicine, including screening, diagnosis, treatment and follow-up. Moreover, the indicators cover a variety of modes of care provision, such as history, physical examination, laboratory study, and medication. Many of the oncology/HIV indicators (42%) are in the "Other Intervention" modality, which includes surgery and radiation therapy.

There are many advantages to a comprehensive quality assessment system. Not only does it cover a broad range of health conditions experienced by the population, but it is also designed to detect underutilization of needed services. In addition, because of its broad scope, it will be difficult for health care organizations to improve their quality scores by focusing their improvement efforts on only a few indicators or clinical areas.

Finally, this system can be effective for both internal and external quality reviews. Sufficient clinical detail exists in the system such that organizations will be able to use the resulting information to improve care, while the simple summary scores that the system generates will be an important tool for health care purchasers and consumers.

ORGANIZATION OF THIS DOCUMENT

The rest of this volume is organized as follows:

- Each chapter summarizes:
 - Results of the literature review for one condition.
 - Provides a table of the staff's recommended indicators based on that review.
 - Indicates the level of scientific evidence supporting each indicator along with the specific relevant citations.
- *Appendix A* provides the summary rating sheets for each condition.
- *Appendix B* shows the changes made to each indicator during the panel process, the reasons for those changes, and the final disposition of each indicator.

REFERENCES

- Brook, R. H., "The RAND/UCLA Appropriateness Method," *Clinical Practice Guideline Development: Methodology Perspectives*, AHCPR Pub. No. 95-0009, Rockville, MD: Public Health Service, 1994.
- Kerr E. A., S. M. Asch, E. G. Hamilton, E. A. McGlynn (eds.), *Quality of Care for Cardiopulmonary Conditions: A Review of the Literature and Quality Indicators*, Santa Monica, CA: RAND, MR-1282-AHRQ, 2000a.
- Kerr E. A., S. M. Asch, E. G. Hamilton, E. A. McGlynn (eds.), *Quality of Care for General Medical Conditions: A Review of the Literature and Quality Indicators*, Santa Monica, CA: RAND, MR-1280-AHRQ, 2000b.
- Kravitz R. L., M. Laouri, J. P. Kahan, P. Guzy, et al., "Validity of Criteria Used for Detecting Underuse of Coronary Revascularization," *JAMA* 274(8):632-638, 1995.
- Malin, J. L., S. M. Asch, E. A. Kerr, E. A. McGlynn. "Evaluating the Quality of Cancer Care: Development of Cancer Quality Indicators for a Global Quality Assessment Tool," *Cancer* 88:701-7, 2000.
- McGlynn E. A., C. Damberg, E. A. Kerr, M. Schuster (eds.), *Quality of Care for Children and Adolescents: A Review of Selected Clinical Conditions and Quality Indicators*, Santa Monica, CA: RAND, MR-1283-HCFA, 2000a.
- McGlynn E. A., E. A. Kerr, C. Damberg, S. M. Asch (eds.), *Quality of Care for Women: A Review of Selected Clinical Conditions and Quality Indicators*, Santa Monica, CA: RAND, MR-1284-HCFA, 2000b.
- Park R. A., Fink A., Brook R. H., Chassin M. R., et al., "Physician Ratings of Appropriate Indications for Six Medical and Surgical Procedures," *AJPH* 76(7):766-772, 1986.
- Schuster M. A., S. M. Asch, E. A. McGlynn, et al., "Development of a Quality of Care Measurement System for Children and Adolescents: Methodological

Considerations and Comparisons With a System for Adult Women," *Archives of Pediatrics and Adolescent Medicine* 151:1085-1092, 1997.

Selby J. V., B. H. Fireman, R. J. Lundstrom, et al., "Variation among Hospitals in Coronary-Angiography Practices and Outcomes after Myocardial Infarction in a Large Health Maintenance Organization," *N Engl J Med* 335:1888-96, 1996.

Shekelle P. G., J. P. Kahan, S. J. Bernstein, et al., "The Reproducibility of a Method to Identify the Overuse and Underuse of Medical Procedures," *N Engl J Med*, 338:1888-1895, 1998b.

Shekelle P. G., M. R. Chassin, R. E. Park, "Assessing the Predictive Validity of the RAND/UCLA Appropriateness Method Criteria for Performing Carotid Endarterectomy," *Int J Technol Assess Health Care* 14(4):