Economic Models of Colorectal Cancer Screening in Average-Risk Adults: Workshop Summary

Michael Pignone, Louise Russell and Judith Wagner, Editors, Institute of Medicine and National Research Council

ISBN: 0-309-54769-5, 310 pages, 8 1/2 x 11, (2005)

This free PDF was downloaded from:
http://www.nap.edu/catalog/11228.html

THE NATIONAL ACADEMIES
Advisers to the Nation on Science, Engineering, and Medicine

Economic Models of Colorectal Cancer Screening in Average-Risk Adults

Workshop Summary

National Cancer Policy Board

Board on Science, Technology, and Economic Policy
Policy and Global Affairs Division

Michael Pignone, Louise Russell and Judith Wagner, *Editors*

INSTITUTE OF MEDICINE *AND*
NATIONAL RESEARCH COUNCIL
OF THE NATIONAL ACADEMIES

THE NATIONAL ACADEMIES PRESS
Washington, D.C.
www.nap.edu

Printed in the United States of America.

THE NATIONAL ACADEMIES
Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Bruce M. Alberts is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Wm. A. Wulf is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Bruce M. Alberts and Dr. Wm. A. Wulf are chair and vice chair, respectively, of the National Research Council.

www.national-academies.org

NATIONAL CANCER POLICY BOARD

JOSEPH SIMONE (*Chair*), Simone Consulting , Dunwoody, GA

ELLEN STOVALL (*Vice Chair*), Executive Director, National Coalition for Cancer Survivorship, Silver Spring, MD

BRUCE W. STILLMAN (*Vice Chair*), Director, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

JILL BARGONETTI, Associate Professor, Department of Biological Sciences, Hunter College, New York, NY

TIMOTHY EBERLEIN, Bixby Professor and Chairman, Department of Surgery, Washington University School of Medicine, St. Louis, MO

KATHY GIUSTI, President, Multiple Myeloma Research Foundation, New Cannaan, CT

KAREN HERSEY, Senior Counsel, Massachusetts Institute of Technology, Cambridge, MA

JIMMIE C. HOLLAND, Chair, Department of Psychiatry and Behavioral Sciences, Memorial Sloan-Kettering Cancer Center, New York, NY

WILLIAM KAELIN, Professor, Harvard Medical School, Boston, MA

WILLIAM McGUIRE, Chief Executive Officer, UnitedHealth Group, Minnetonka, Minnesota

JOHN MENDELSOHN, President, M. D. Anderson Cancer Center, University of Texas, Houston, TX

KATHLEEN HARDIN MOONEY, Professor, University of Utah College of Nursing, Salt Lake City, UT

PATRICIA NOLAN, Director, Rhode Island Department of Health, Providence, RI

DAVID PARKINSON, Oncology Therapeutic Area Head, Amgen, Inc. Thousand Oaks, CA

JOHN POTTER, Senior Vice President and Director, Cancer Prevention Research Program, Fred Hutdchinson Cancer Research Center, Seattle, WA

LOUISE RUSSELL, Professor, Rutgers University, New Brunswick, NJ

THOMAS J. SMITH, Professor, Virginia Commonwealth University, Richmond, VA

ROBERT C. YOUNG, Past-President, American Cancer Society and President, the Fox Chase Cancer Center, Philadelphia, PA

Study Staff

JUDITH L. WAGNER, Scholar in Residence, Institute of Medicine

ELIZABETH J. BROWN, Research Assistant

ANIKE L. JOHNSON, Administrative Assistant

MARY ANN F. PRYOR, Project Assistant

LOUISE B. RUSSELL, Research Professor, Institute for Health, Health Care Policy and Aging Research, Rutgers University, Workshop Chair

MICHAEL P. PIGNONE, Assistant Professor of Medicine, University of North Carolina

v

REVIEWERS

All presenters at the workshop have reviewed and approved their respective sections of this report for accuracy. In addition, this workshop summary has been reviewed in draft form by independent reviewers chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the National Research Council's Report Review Committee. The purpose of this independent review is to provide candid and critical comments that will assist the Institute of Medicine (IOM) in making the published workshop summary as sound as possible and to ensure that the workshop summary meets institutional standards. The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process.

The National Cancer Policy Board and IOM thank the following individuals for their participation in the review process:

John Inadomi, University of Michigan Medical Center and VA Ann Arbor Health Care System
David Lieberman, Oregon Health Sciences University
Diana Petitti, Kaiser Permanente of Southern California
David Ransohof, University of North Carolina at Chapel Hill

The review of the report was overseen by **Clyde J. Behney, Deputy Executive Director, IOM,** who was responsible for making certain that an independent examination of this report was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of this report rests entirely with the editors and the institution.

TABLE OF CONTENTS

TABLES AND FIGURES

TABLES

FIGURES

viii

Workshop Summary

INTRODUCTION AND BACKGROUND

Colorectal cancer (CRC) is the second leading cause of death from cancer in the United States (Edwards et al., 2002). Research has shown that screening adults for early cancers or their precursor lesions, followed by appropriate therapy and continued surveillance, can reduce CRC incidence and mortality (Curry, 2003). A general consensus has emerged that periodic screening of adults over age 50 is a valuable preventive intervention and today most health plans cover CRC screening (United States General Accounting Office, 2004). Yet, there is continued uncertainty about the specific screening strategies that should be offered to individuals who are at average risk for CRC.

There are two reasons for the prevailing uncertainty about what screening strategies make sense for these average-risk adults. First, the number of potential screening strategies is large, encompassing not only the choice of technology (or technologies) but also decisions about the age at which screening should begin, the frequency with which it should occur, and the age at which routine screening should end. Several medical technologies are available to detect early cancers or benign adenomas, the polyps that precede most colorectal cancers. Those technologies vary widely both in cost and detection capabilities. The list includes flexible sigmoidoscopy, colonoscopy, barium enema x-ray, and fecal occult blood tests. The choices are growing, too. New technologies, including imaging and molecular markers, are currently under development. Their entry will expand the range of alternative screening strategies even further.

A second factor that makes it difficult to settle on a specific strategy is that much is unknown about the natural history of colorectal cancer—how fast or slowly it develops, how frequently it arises from pre-existing benign adenomas, and how long those adenomas remain in a benign but detectable state before they convert to cancer. Although new information about these questions has emerged in recent years, it is indirect because, once they are detected, cancers or adenomas are virtually never left behind to grow and be observed. The effectiveness and cost of any screening strategy depend on the details of natural history and as long as those details remain unknown, it is impossible to be sure that one strategy is unequivocally better than another in the absence of a head-to-head trial comparing different strategies. Such a trial is unlikely to be performed because the cost and duration would be prohibitive.

Economic models of CRC screening offer a means for addressing questions about how to screen for CRC. Beginning with the work of David Eddy in the late 1970s (Eddy, 1980), many academic and government researchers have built computer models to describe the natural history of CRC and analyze the costs and effects of altering that history with selected screening strategies (Eddy et al., 1987; Frazier et al., 2000; Glick et al., 1998; Joseph et al., 1988; Khandker et al., 2000; Ladabaum et al., 2004b; Lieberman, 1995; Loeve et al., 1999, 2000; Neilson and Whynes, 1995; Ness et al., 2000; Sonnenberg and Delco, 2002; Sonnenberg et al., 2000; Vijan et al., 2001; Wagner et al., 1991, 1996; Whynes, 2004). The purpose of such models is to help decision makers evaluate which strategies to pay for, recommend, adopt, or use. As the field of cost-effectiveness analysis (CEA) in medicine advances (Gold et al., 1996), and as new evidence on the natural history of CRC emerges, the models have improved. But they have not been able to resolve the uncertainty about the comparative performance of different CRC screening strategies. Rather, they continue to disagree about how alternative strategies stack up against one another in their health effects and costs (Curry, 2003; Pignone et al., 2002).

1

Public health policy makers increasingly rely on CEAs to help them sift through the many choices confronting them. When different CEA models give different answers to the same question, confidence in their usefulness may suffer, since it is unclear to what extent the disagreement arises from uncertainty about the underlying evidence, which affects all decision making approaches, or from the modeling methods used by different modelers. Understanding the reasons for differences among models is therefore an important first step in building the public's confidence that CEA can provide objective and informative insights into the consequences of health policy choices.

The Institute of Medicine's (IOM's) National Cancer Policy Board (NCPB) convened the workshop, *"Economic Models of Colorectal Cancer Screening in Average-Risk Adults"* on January 26-27, 2004, to explore the reasons for differences among leading CEA models of CRC screening. Participants discussed the results of a collaborative pre-workshop exercise undertaken by five research teams that have developed and maintained comprehensive models of CRC screening in average-risk adults. The purpose of the exercise was to provide workshop participants with insights into each model's structure and assumptions and possible explanations for differences in their published analyses. Workshop participants also examined the current state of knowledge on key inputs to the models with a view toward identifying areas where further research may be warranted.

In keeping with the purpose of IOM workshops, this summary of its proceedings presents the individual perspectives and research of people who made presentations at the workshop and of many other experts who participated. This summary does not contain consensus recommendations, nor does it represent a consensus opinion of the IOM's NCPB. Nor is it intended as a guide for conducting or using cost-effectiveness analyses in CRC screening decisions.

It is particularly important to recognize that the purpose of the workshop was *not* to consider the relative merits of different strategies for CRC screening, or to suggest which CRC screening strategy is best. It was solely to consider the commonalities and differences among the CEA models bearing on the subject. The demand for more certain guidance from models by those who recommend or pay for screening strategies, while clearly a motivating force behind the workshop, was not its focus. More certain guidance may result in the future as modelers continue to grapple with and explain the differences in their findings.

THE COLLABORATIVE MODELING EXERCISE

Origin of the Exercise

The idea for collaboration among research teams that maintain published models of CRC screening grew out of a recent review by Michael Pignone and colleagues for the U.S. Preventive Health Services Task Force (Pignone et al., 2002). They systematically reviewed seven published CEAs of periodic CRC screening in average-risk adults. That review identified several aspects of model structure and underlying assumptions which, taken together, might account for most of the differences in cost-effectiveness rankings of CRC screening strategies. However, each model involves dozens of assumptions, and the reviewers concluded that the published reports provided insufficient information to determine which assumptions or aspects of model design were most important in explaining differences in conclusions across models.

The goal of the collaborative pre-workshop exercise was to shed light on the degree to which difference across models could be reduced by standardizing the values of key input parameters, or assumptions, across models. Any residual variation in model outcomes would be the result of differences either in parameters that remained unstandardized or in the structure of the models themselves. Secondary objectives were to demonstrate the benefit of collaborative interactions among modelers and to ascertain the research resources (time and money) required to mount such exercises.

General Approach

Five research teams with published CEAs of colorectal cancer screening agreed to participate in a comparative modeling exercise to further explore the reasons for disparate cost-effectiveness findings. Each of the models can track (via computer) a hypothetical cohort of average-risk Americans, beginning at age 50, over their remaining lifetimes and can estimate the number of years of life lived and the medical costs incurred by the members of that cohort.[1] The participating research teams were:

- The Harvard Model (Frazier et al., 2000), led by Karen Kuntz, Ph.D.;
- The Ladabaum Model (Ladabaum et al., 2004a; Song et al., 2004), led by Uri Ladabaum, M.D.;
- The Miscan Model (Loeve et al., 1999, 2000), led by Marjolein van Ballegooijen, M.D.;
- The Vanderbilt Model (Ness et al., 2000) led by Reid Ness, M.D.; and
- The Vijan Model (Vijan et al., 2001), led by Sandeep Vijan, M.D.

At the workshop, each team leader described essential features of the model's structure and assumptions. (See the appendixes with speakers' presentations.) The teams further agreed to provide cost-effectiveness results for a set of five specific screening strategies across 10 different combinations of assumptions, starting with the assumptions in their original models.

The Screening Strategies

All the strategies included in the pre-workshop exercise envisioned periodic screening of all average-risk Americans beginning at age 50 and ending at age 80. The five selected strategies were:

1) **F/S**: Annual fecal occult blood testing in combination with a flexible sigmoidoscopy every five years;
2) **S**: Sigmoidoscopy every five years;
3) **R**: A prototype radiology procedure every five years, with specific test characteristics and costs;
4) **C**: Colonoscopy every 10 years; and
5) **F**: Annual fecal occult blood testing.

These strategies were selected not for any posited superiority over other CRC screening approaches, but for the frequency with which they are advocated by practitioners today. Some of them represent strategies that have been recommended by professional groups (Smith et al., 2004; U.S. Preventive Services Task Force, 2002; Winawer et al., 2003). They also represent a wide range of procedure cost and test accuracy.

---

[1] Some of models can track all age cohorts of adults over a long period of time as well as specific age cohorts.

The prototype radiology strategy differed from the others by virtue of being defined by specific assumptions about costs and test performance. That route was necessary because some research teams had not investigated CRC screening with radiological technologies and therefore had no original assumptions at the ready. Moreover, an emerging imaging technique—virtual colonoscopy—may eventually join a much older radiology procedure—double-contrast barium enema (DCBE)—as an entry in the mix of available screening technologies. (Cotton et al., 2004; Pickhardt et al., 2003; Ransohoff, 2004). The assumptions specified for the prototype strategy represent an optimistic mix of cost and test performance characteristics based on the old and new radiology procedures.

The Standard Assumptions

The pre-workshop exercise specified standard assumptions in each of four groups listed below:

1) follow-up and periodic surveillance regimens—the assumptions that modelers make about how the health care system responds to a positive screening test, both in the short term (diagnostic follow-up) and after removal of a pre-cancerous adenoma (surveillance);

2) test performance characteristics—the sensitivity, specificity, and medical risk of tests for screening, follow-up, and periodic surveillance after treatment;

3) medical costs—the costs of screening, follow-up, and surveillance, as well as the costs of treating colorectal cancer at various stages; and

4) compliance—expected levels of adherence to the screening, follow-up, and surveillance strategies under evaluation.

The standardized assumptions in each of these groups are shown in Table 1.

TABLE 1. Standardized Assumptions for Pre-workshop Collaborative Exercise

COSTS	
Fecal occult blood est	$10
Colonoscopy-diagnostic	$625
Colonoscopy with polypectomy	$900
Pathology per polyp	$65
Sigmoidoscopy-screening	$200
Sigmoidoscopy with polypectomy	N/A
Sigmoidoscopy with biopsy	$375
Prototype tadiology procedure	$200
Lifetime CRC treatment cost	
Local	$24,000
Regional	$31,000
Distant	$40,000
Cost of treating perforation	$24,000

TABLE 1 Continued

TEST PERFORMANCE	
Sigmoidoscopy	
Reach (percent of polyps)	50 percent
Sensitivity for polyps	85 percent
Sensitivity for cancer	95 percent
Specificity	100 percent
Fecal occult blood test-not rehydrated	
Sensitivity for polyps	10 percent
Sensitivity for cancer	40 percent
Specificity	97 percent
Colonoscopy	
Sensitivity for polyps	85 percent
Sensitivity for cancer	95 percent
Specificity	100 percent
Prototype radiology procedure	
Sensitivity for polyps	70 percent
Sensitivity for cancer	80 percent
Specificity	90 percent
Complications	
Colonoscopy major complications (perforation)	0.10 percent
Colonooscopy mortality rate	0.01 percent
Sigmoidoscopy major complication	0 percent
Sigmoidoscopy mortality rate	0 percent
Prototype radiology	0 percent
FOLLOW-UP	
Fecal occult blood test	Assume all positive fecal occult blood tests are followed by colonoscopy with polypectomy if true positive, or diagnostic colonoscopy if false positive
Sigmoidoscopy	a) Assume all positive screens are followed by colonoscopy with polypectomy if true positive, or diagnostic colonoscopy if false positive b) Assume positive screen involves no biopsy or polypectomy—cost is for screening sigmoidoscopy
Colonoscopy	Assume positive screen involves polypectomy with biopsy
Prototype radiology	Assume all positive screens are followed by colonoscopy with polypectomy if confirmed as true positive, or diagnostic colonoscopy if false positive

TABLE 1 Continued

SURVEILLANCE	All individuals with adenomatous polyps get surveillance with colonoscopy every 5 years, beginning with fifth year post-polypectomy. Continued until 80 years of age or death
COMPLIANCE	100 percent with all aspects of strategy (screen, follow-up and surveillance)

A small number of basic assumptions, such as the discount rate, were also specified to remove possible sources of variation among models deriving from technical details (see Table 2).

Each research team first produced results with its own original assumptions, as shown in Table 3.[2] Then they produced results in successive runs when assumptions in one group at a time were assigned standardized values, leaving the rest at their original values. They generated a third set of results for a series of runs when one group of assumptions was left at its original values while the rest of the groups were standardized. A final run produced estimates when all assumptions in the exercise were standardized.

The standardized assumptions were *not* selected with the goal of specifying "correct" values. For the most part they were selected to strike a compromise among the five research teams' original assumptions. However, some values were set to accommodate the least specific model in order to avoid the need for extensive reprogramming. For example, standardized compliance was set at 100 percent. Although an abundance of evidence suggests compliance is far less than perfect, it would have been time-consuming or impossible for all of the research teams to reconfigure their models to accommodate more realistic assumptions. This somewhat opportunistic standardization process underscores the danger of interpreting the standardized results as endorsing any specific colorectal cancer screening strategy, especially because the effectiveness of some strategies is bound to be more heavily dependent on high rates of compliance than others.

TABLE 2. Basic Assumptions

Population size (at 50 years of age):	100,000
Population demographics:	Average-risk individuals, U.S. population, both sexes and all races
Discount rate:	3 percent per year
Quality adjustments for all health states short of death:	None
Type of output:	Cohort model, followed from age 50 through age 85

---

[2] Researchers were given the choice of using the same assumptions as those made in their published papers or using other assumptions if more recent work had led them to new assumptions in the current versions of their models.

TABLE 3. Summary of Assumptions in Five Economic Models of CRC Screening

	Harvard	Laudabaum	Miscan	Vanderbilt	Vijan
TEST PERFORMANCE					
FOBT					
Rehydrated?	No	No	NA	No	No
Sensitivity-small adenoma	0.1	0.08	0.02	0.05 (<=10 mm)	0.05
Sensitivity-large adenoma	0.1	0.1	0.05	0.11	0.05
Sensitivity-high-risk adenoma	0.1	0.1	-	0.11	0.05
Sensitivity-early CA	0.33	0.4	0.6	0.13	0.3
Sensitivity-regional CA	0.33	0.4	0.6	0.13	0.5
Sensitivity-late stage CA	0.33	0.4[a]	0.6	0.13	Not modeled[c]
Specificity	0.97	0.92	0.98	0.95	0.975
Test performance on each screen independent?	Yes	Yes	Yes	Yes	Yes
Medical risk of procedure	None	None	None	None	None
Each screen Independent?	Yes	Yes	Yes	Yes	Yes
FSIG					
Average percentage of lesions reached	About 50[d]	50	64		55 percent
Sensitivity-small adenoma	0.85	0.7	0.75 for <=5 mm, 0.85 for 6-10 mm	0.75 (<5 mm); 0.80 (5-10mm)	0.85
Sensitivity-large adenoma	0.95	0.8	0.95	0.85	0.95
Sensitivity-high-risk adenoma	0.95	0.8	-	0.85	0.95
Sensitivity-early CA	0.95	0.9	0.95	0.95	0.95
Sensitivity-regional CA	0.95	0.9	0.95	0.95	0.95
Sensitivity-late stage CA	0.95	0.9[a]	0.95	0.95	0.95
Specificity	1	0.95	0.92-hyperplstic polyp = FP	1	1
Risk of perforation	Not modeled explicitly—cost of complications embedded in procedure cost	0.0001[b]	0	0.0001	0
Risk of bleeding			0	0	0
Risk of death—complications	0.0000014	0.00001 (0.0001× 0.10)	0	0.00002	0

7

continued

TABLE 3 Continued

	Harvard	Laudabaum	Miscan	Vanderbilt	Vijan
Each screen independent?	Yes	Yes	Yes	Yes	Yes
Other?				Non-adenoma found in random 20 percent of exams—removed with additional risk and cost	
CSCPY					
Sensitivity for small adenomas	0.85	0.85	0.80 for <=5mm, 0.85 for 6-10mm	0.75 (<5 mm); 0.80 (5-10mm)	0.85
Sensitivity for large adenomas	0.95	0.9	0.95	0.85	0.95
Sensitivity for high-risk adenomas	0.95	0.9	-	0.85	0.95
Sensitivity for early CA	0.95	0.95	0.95	0.95	0.95
Sensitivity for regional CA	0.95	0.95	0.95	0.95	0.95
Sensitivity for late stage CA	0.95	0.95[a]	0.95	0.95	0.95
Specificity	1	1	1	1	1
Risk of perforation	Not modeled explicitly-cost of complications embedded in cost of procedure	0.001[b]	0.002	0.001	0.001
Risk of bleeding			-	0.009	0
Risk of death from complications	0.00005	0.0001(0.001 × 0.10)	0.0001	0.0002	0.075
Each screen independent?	Yes	Yes	Yes	Yes	Yes
Other?				Non-adenoma found in random 26 percent of exams- removed with additional risk and cost	
FOLLOW-UP ASSUMPTIONS					
FOBT					
Does positive FOBT result in colonoscopy?	Yes	Yes	Yes	Yes	Yes
FSIG					
Is polyp found on screen removed in same procedure?	Yes	No	No	No	No
Is polyp found on screen biopsied?	NA	No, but CRC biopsied	No	No	Yes
Is a polyp found on screen referred to colonoscopy w/o biopsy or removal?	NA	Yes	Yes	Yes	No

continued

8

TABLE 3 Continued

	Harvard	Laudabaum	Miscan	Vanderbilt	Vijan
Is a polyp found on screen referred to colonoscopy after biopsy or removal?	Yes	No	No	No	Yes
SURVEILLANCE ASSUMPTIONS					
Is surveillance contingent on size or type of polyp? (if yes, explain)	Yes High-risk polyps only	No	Yes Depends on size and number of polyps	Yes Adenomas only	Yes Only for those with >=1cm or multiple polyps
Describe colonoscopic surveillance schedule	Every 3 years	Every 5 years until age 80	5 years if <3 small adenomas 3 years if >=3 small adenomas 3 years if >=1 large adenomas 5 years after negative surveillance No age bounds	Colonoscopy every 5 years	Colonoscopy every 5 years
COST ASSUMPTIONS					
FOBT (unit cost)	$38	$20	$4.50	$4	$18.40
FSIG (unit cost)	$279	$290	$100	$401	$389
FSIG with biopsy (unit cost)	Embedded in polypectomy cost	$440		NA	$397
FSIG with polyp removal	$528(279+249)	NA		NA	Not modeled
CSCPY-diagnostic	$1,012	$820	$650	$681	$653
CSCPY with polyp removal	$1,065(1012+553)	$1,200	$750	$808	$737
Pathology	Embedded in polypectomy cost	Included in above	Included in CSCPY with polypectomy	$181	$95
Cost of treating complications	Embedded in colonoscopy cost	$26,000	$30,000	Hemorrhage:$4,360; perforation/death:$13,000	$13,361
Cost of treating CRC-local (lifetime to age 85—discounted 3 percent)	$21,941	$46,000	$26,800		$34,252
Initial cost	$15,000			$17,843 (first 6 months)	
Continuing cost	$392/year		$2,100/year	$473/yr (prorated)	
Final cost	$18,014			$18,589 (last 6 months)	
Cost of treating CRC-regional (lifetime to age 85—discounted 3 percent)	$43,623	$68,000	$26,800		$47,227
Initial cost	$17,920			$20,517 (first 6 months)	
Continuing cost	$1793/year		$2,100/year	$2,161/year (prorated)	
Cost of treating CRC-distant (lifetime to age 85—discounted 3 percent)	$58,231	$71,000			$41,602

continued

9

TABLE 3 Continued

	Harvard	Laudabaum	Miscan	Vanderbilt	Vijan
Initial cost	$29,235		$26,800	$23,449 (first 6 months)	
Continuing cost	$19,562/year		$2,100/year	$23,578/year (prorated)	
Final cost	$15,531			$18,589 (last 6 months)	
COMPLIANCE ASSUMPTIONS					
Does compliance with screening vary by type of patient? (if so explain)	No	No	No	Yes	No
Does compliance with screening change for individual patient over time?	No	No	No	No	No
Is patient's compliance at any time dependent on prior compliance?	No	No	Depends on past compliance	Yes	Yes
Does compliance with screening vary with screening technology?	Yes	No	Yes	Yes	No
Compliance with FOBT	0.6				
Compliance with FSIG	0.6				
Compliance with CSCPY	0.6				
Compliance with CSCPY used in surveillance or as a F/U test	0.8				
Describe approach to modeling compliance	Persons have a "p" perecent chance of complying in a screening year	Compliance is all-or-nothing	Colonoscopy—first: 30 percent; subsequent: if previously Attended, 63.2 percent; if previously not attended, 15.8 percent Radiology: 40 percent, 72.7 percent, 18.2 percent FSIG: 40 percent, 72.7 percent, 18.2 percent FSIG/FOBT: 35 percent, 68.3 percent, 17.1 percent FOBT: 60 percent 85.7 percent, 21.4 percent	Compliance varies with frequency of screening and screening modality. FOBT: every 1 years–35 percent every 2 years–14 percent every 5 years–11 percent one time–10 percent never–30 percent Endoscopy/Radiology every 5 years–53 percent one time–17 percent never–30 percent Surveillance Endoscopy 80 percent (80-100)	

NOTES: All assumptions are based on current versions (as of December 20, 2003) of researchers' models, except for Harvard model, which followed assumptions published in Frazier et al. (2000). *a.* not relevant-distant cancers assumed to present with symptoms within the year; *b.* includes all major complications; *c.* assumed all late-stage cancers are symptomatic; *d.* distal and rectal lesions, age-and sex-specific

10

Note also that assumptions about the natural history of colorectal cancer screening differ across models, but standardizing those assumptions is especially difficult to do and was not attempted. Natural history assumptions—the prevalence and incidence of adenomas and other benign polyps, how fast adenomas progress to cancer, what proportion of cancers are preceded by benign adenomas, and how fast cancers progress from early to late stages, and life-expectancy of the population with and without colorectal cancer—are interrelated with one another. They can be specified at various levels of detail, by age, sex and race, or other risk factors, as well as by location of the lesion in the colon and by the existence of past or concurrent adenomas. Some models can incorporate very detailed natural history assumptions, whereas others cannot. Additionally, model structures vary in the kind of natural history inputs required. For example, some models require data on the monthly or annual probability that an adenoma will progress to early cancer, whereas others require estimates of the number of years of growth required before an adenoma makes the transition to colorectal cancer. Because of these difficulties, the research teams agreed that the comparative modeling exercise should not attempt to standardize assumptions regarding natural history. Instead, they agreed to provide some intermediate results: the number of adenomas or polyps detected, deaths from CRC, and total mortality at each age between 50 and 85 in the absence of screening. Those results would allow an indirect comparison of natural history assumptions across the models.[3]

Specification of Model Outputs

For every model run, the research teams provided the coordinators of the exercise[4] with estimates of the total number of years of life lived and total medical costs incurred by a population of 100,000 average-risk 50-year-old adults from age 50 until death or age 85, whichever comes first.[5] These outputs were reported both as simple totals and in terms of their net present value (NPV) at the starting age (age 50).[6]

The cost-effectiveness of any screening strategy compared with any other strategy or with no screening, may be calculated from those outputs. For example, the cost-effectiveness of a strategy compared with no screening at all is as follows:

$$CE = \frac{\textit{(Lifetime cost with strategy – Lifetime cost with no screening)}}{\textit{(Years lived with strategy – Years lived with no screening)}}$$

If both numerator and denominator are positive, then the C/E ratio represents the extra costs required to achieve each extra year of life. If the numerator of the ratio is negative, while the denominator is positive, then the strategy saves both costs and lives and is unequivocally superior to doing nothing.

---

[3] The intermediate results were not presented at the workshop and are therefore not discussed in this summary.

[4] Four workshop participants, Martin Brown, Louise Russell, Michael Pignone, and Judith Wagner, led the development of the pre-workshop exercise and coordinated the analysis of its results. Michael Pignone presented the analysis at the Workshop (See his presentation in Appendix I.)

[5] Screening programs lasted 30 years, but the reporting period continued for 35 years.

[6] All comparisons using NPV applied an annual discount rate of 3 percent (Gold et al., 1996).

TABLE 4. Assumption Settings for Pre-workshop Exercise

Run Number	Cost	Test Performance	Follow-Up and Surveillance	Compliance
1	Orig	Orig	Orig	Orig
2	Std	Orig	Orig	Orig
3	Orig	Std	Orig	Orig
4	Orig	Orig	Std	Orig
5	Orig	Orig	Orig	Std
6	Std	Std	Std	Std
7	Std	Std	Std	Orig
8	Std	Std	Orig	Std
9	Std	Orig	Std	Std
10	Orig	Std	Std	Std

NOTES: "Std" = all assumptions in the group are standardized (see Table 2); "Orig" = all assumptions in the group are set at their original values (see Table 3).

The Comparisons

The five research teams were asked to report results for the baseline—no screening—as well as for 10 runs for each of the five screening strategies, 50 runs in all, as noted in Table 4. Each team ran its model 52 times (twice for the no-screening strategy,[7] and 10 times for each of the 5 screening strategies). Thus, the research teams submitted a total of 260 separate computer runs for analysis by the coordinators

Two runs represent the extremes of the standardization spectrum. Run number 1 produced results for the model's original assumptions in all four areas—follow-up, test performance, cost, and compliance. Run number 6 showed the results when all assumptions were set to their standardized values. All other model runs involved combinations of original and standardized assumptions.

Results

Baseline Estimates (No Screening)

The research teams estimated the number of years of life lived (life expectancy) by an average 50-year-old and lifetime CRC-related costs per person, when no screening program was in effect and all assumptions were set to each team's original values (Table 5). Any differences among models in those estimates would reflect variations either in model structure or in assumptions about age-specific mortality in the U.S. population, age- and stage-specific incidence of colorectal cancer, and costs of treating colorectal cancer by age and stage.

The research teams reported a range of estimates of years of life lived. The average life expectancy in the model with the highest predicted value was about 2.25 years or 1.1 times longer than in the model with the lowest value. Two models predicted almost identical life expectancies of 25 years; three predicted identical life expectancies of 27 years. In reviewing these results at the workshop, several researchers suggested that the

---

[7] The no-screening strategy required Runs #1 and #2, because standardizing costs of treating colorectal cancer (as in Run # 2) would change model outcomes even without screening. All other runs involve changes in assumptions that would occur only under a screening regimen.

TABLE 5. Predicted Years of Life Lived and Lifetime per-Capita CRC Costs No Screening: Original Model Assumptions

Model	Years of Life Lived	Rank	Lifetime Costs	Rank
Harvard	25.12	2	$1,811	3
Ladabaum	27.23	3	$2,994	4
Miscan	27.29	4	$2,283	5
Vanderbilt	27.39	5	$1,792	2
Vijan	25.07	1	$1,665	1

NOTE: Results are not discounted to net present value.

differences were due to the use of mortality statistics from different years. Sandeep Vijan and Karen Kuntz remarked that assumptions about life expectancy in their models were based on older life tables. Mortality rates have decreased substantially in the last decade, especially in older age groups.

The variation among models reported by the research teams in estimated lifetime costs was larger than the variation in effects, with the highest estimate about 1.8 times higher than the lowest. Those disparities reflect the models' very different assumptions about and approaches to estimating the cost of treating colorectal cancer. When treatment costs were standardized to the values shown in Table 1, the range of estimated costs diminished substantially to a ratio of 1.2 between the highest and lowest values. Some participants posited that differences in assumptions about cancer incidence probably account for the remaining variation in colorectal cancer costs.

Screening Estimates Under Original Assumptions

Differences among the five models in estimates of the effect of screening under each team's original assumptions were presented by Michael Pignone and discussed by the research teams and other participants.

Comparing screening with no screening. Figure 1 shows the net increases in years of life lived and lifetime costs (discounted to their NPV), compared with no screening under the full set of original assumptions adopted by each research team. The research teams reported wide variation in ratio terms for each of these two components of cost-effectiveness. For example, the NPV of lifetime cost reported for a screening program of flexible sigmoidoscopy every 5 years ranged from $224 per person (Miscan) to $1,159 per person (Vanderbilt), a five-fold difference between the two. The predicted gains in life expectancy from screening are less varied than for costs, but still high. For example, the net present value of life-years gained from flexible sigmoidoscopy ranged from 2,723 per 100,000 50-year-olds (Miscan) to 4,265 (Vanderbilt), a ratio between the highest and lowest of about 1.6.

The research teams reported that the most effective strategy differed across the models. Two models predicted that F/S gains the most years of life for the population, two models predicted that R would be most effective, and one model predicted C is the most effective. The least costly strategy also differed across models. Two models predicted that S is the least costly strategy, two that F is least costly, and one that R is least costly[8]

---

[8] In discussion following the presentation of the results, Louise Russell and others pointed out that the analysis focused on differences across models in their single best estimates of effects and costs, whereas the research teams have acknowledged and reported on the range of uncertainty surrounding their estimates in their research papers. Had uncertainty been modeled in this exercise, the range of reported results might have overlapped.
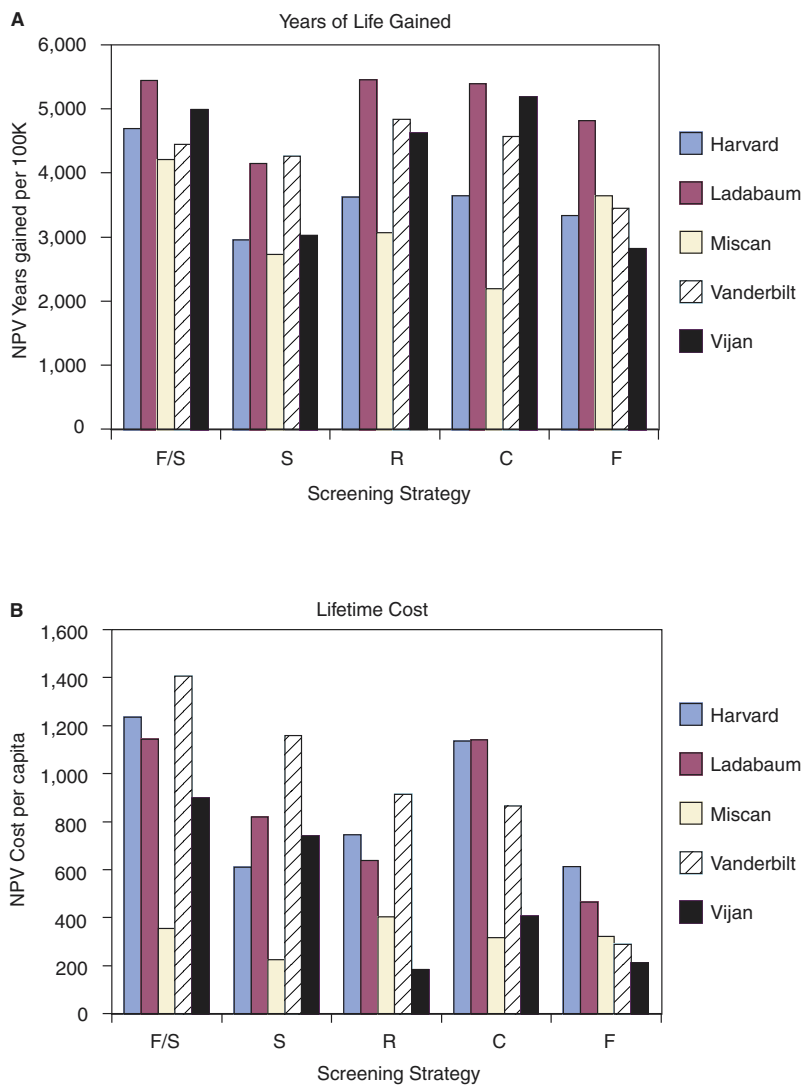
A

Years of Life Gained

B

Lifetime Cost

FIGURE 1. Years of life gained and liftetime costs of screening: original assumptions.
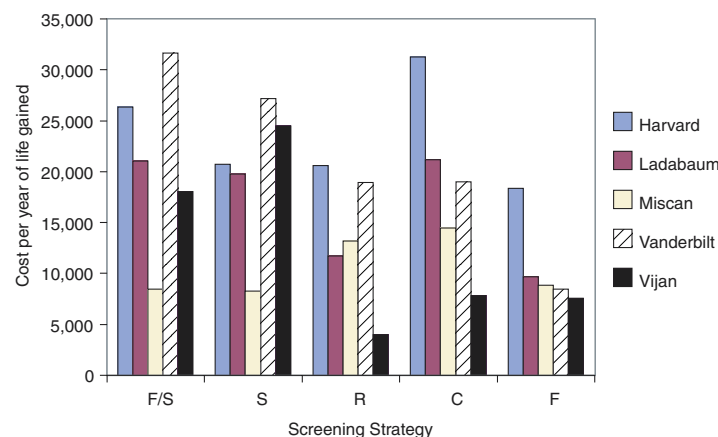
FIGURE 2. Cost effectiveness of screening: original assumptions.

As a result, estimates of the cost-effectiveness ratio also varied across the five models, in some cases by a five-fold difference between the highest cost-effectiveness ratio and the lowest (Figure 2). Despite that variation, Michael Pignone pointed out, all the models show that all of the strategies meet common benchmarks of cost-effectiveness. Every research team estimated that, when compared with no screening, colorectal cancer screening could deliver an additional year of life for a cost of less than $40,000, regardless of which strategy is adopted.

Comparing strategies with one another. The goal of cost-effectiveness analysis is to compare alternative strategies with one another (Gold et al., 1996). The disparities among CRC models in such comparisons prompted the Workshop to begin with. So, participants reviewed the performance of strategies with each other as reported by the research teams.

The first step in making such comparisons is to rule out any screening strategy that is both less effective and more costly than at least one other. Strategies ruled out at this stage are referred to as "strongly dominated." The second step is more subtle. It requires ruling out any strategy whose gains in life expectancy, compared with the next most effective strategy, come at an incremental cost that is higher than the incremental cost of achieving gains at least as great through still another strategy. Strategies ruled out at this stage are referred to as "weakly dominated." Any strategies surviving this two-step elimination process present a true trade-off between successively higher costs and greater health benefits. Louise Russell reminded participants, however, that the process is based on point estimates, which are subject to uncertainty. All research groups have routinely assessed the effect of uncertainty on those estimates. Had the exercise included such analyses, it might have found that some strategies that were ruled out were essentially equivalent to those ruled in.

Once the strategies surviving the two rule-out tests are identified, their incremental cost-effectiveness ratios can be calculated by sorting them into ascending order of effectiveness, measuring the differences in both cost and years of life gained compared with the next most effective strategy (or with no screening for the least effective strategy), and calculating the cost-effectiveness ratio. Michael Pignone summarized the results.

TABLE 6. Incremental Cost-Effectiveness Ratios of Five CRC Screening Strategies: Original Assumptions

	Harvard	Ladabaum	Miscan	Vanderbilt	Vijan
F/S	$45,976	SD	$8,848	SD	SD
S	WD	SD	$8,230	SD	SD
R	WD	$27,069	SD	$44,936	$3,980
C	WD	SD	SD	WD	$38,854
F	$18,347	$9,631	WD	$8,409	SD

NOTES: F/S = annual fecal occult blood test; sigmoidoscopy every 5 years; S = sigmoidoscopy every 5 years; R = proto-type radiology procedure every 5 years; C = colonoscopy every 10 years; F = annual fecal occult blood test; WD = strategy is weakly dominated by at least one other strategy; SD = strategy is strongly dominated by at least one other strategy.
SOURCE: M. Pignone Workshop Presentation (Appendix I).

Across the five models, the surviving strategies differed substantially and the incremental cost-effectiveness ratios of those strategies also differed (Table 6). Thus, according to Pignone, under their original assumptions, the five research teams would present very different options to policy makers.

Estimates of Screening Under Standardized Assumptions

Michael Pignone presented the effect of standardizing all of the assumptions in the four groups together on differences among models.

Comparing screening with no screening. Under the full set of standardized assumptions, the two components of the cost-effectiveness ratio still varied across models. Sometimes, but not always, by less in ratio terms than when the models used their original assumptions (Figure 3). Differences across models in predicted per capita lifetime costs were greatest for strategy S, where they ranged from $718 per person (Miscan) to $1,436 per person (Vanderbilt), a two-fold difference between the two. (Recall that the difference was five-fold under the original assumptions.)

Differences across models in years of life gained from screening did not change in a systematic way after standardization. The range of variation grew modestly for two strategies and declined for the other three. The NPV of life-years gained from strategy S ranged from 3,470 per 100,000 50-year-olds (Vijan) to 6,954 (Vanderbilt), a ratio of 2.0 between the highest and lowest, compared with a ratio of 1.6 under the teams' original assumptions. The two strategies involving sigmoidoscopy seemed to resist convergence in predicted years of life gained more than other strategies.

Standardization of assumptions did result in agreement across models on the most effective and least costly strategies. All of the research teams estimated that F/S gains the most years of life and all found F to be the least costly strategy.

The cost-effectiveness ratio for each strategy continued to vary across the five models, but the range of difference as measured by the ratio of the highest to lowest narrowed with full standardization (Figure 4).With all tested assumptions standardized, the cost-effectiveness ratio varied across models by a factor of 1.5 to 2.0 for every strategy.
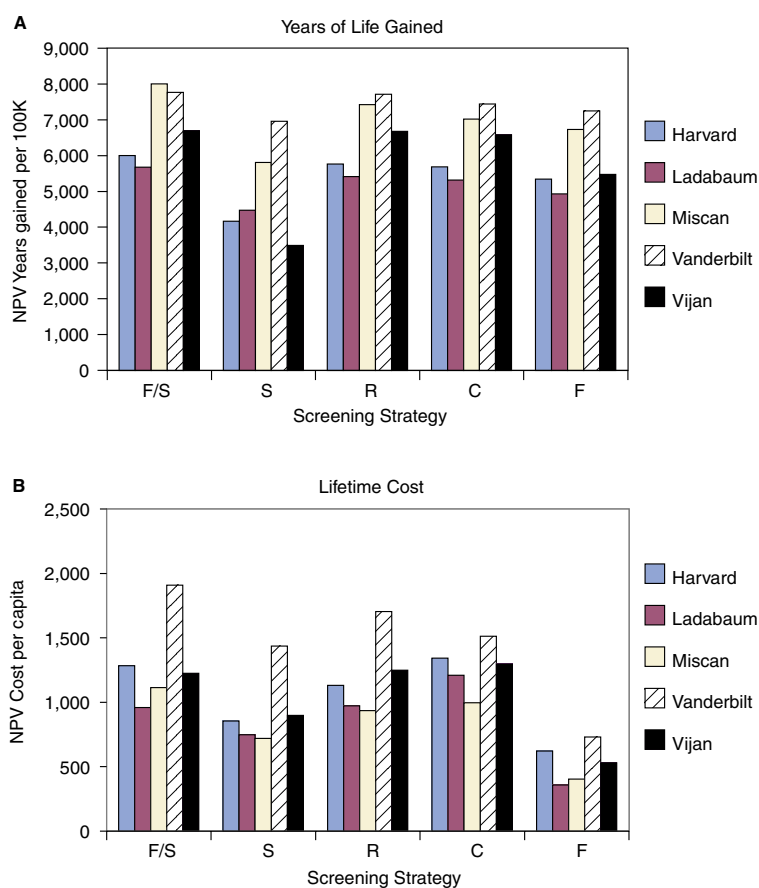
FIGURE 3. Years of life gained and lifetime costs of screening: standardized assumptions.
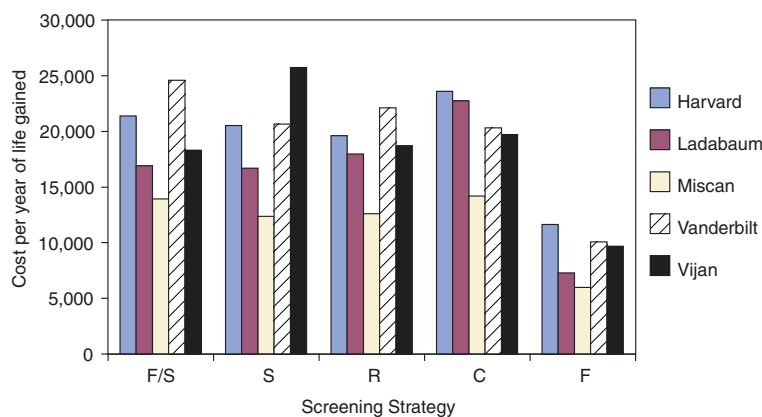
FIGURE 4. Cost-effectiveness of screening: standardized assumptions.

TABLE 7. Incremental Cost-Effectiveness Ratios of Five CRC Screening Strategies: Standardized Assumptions

	Harvard	Ladabaum	Miscan	Vanderbilt	Vijan
F/S	$99,997	$79,920	$55,828	$355,647	$56,969
S	SD	SD	SD	SD	SD
R	WD	SD	WD	$209,906	SD
C	SD	SD	SD	SD	SD
F	$11,632	$7,232	$5,980	$10,073	$9,676

NOTES: F/S = annual fecal occult blood test; sigmoidoscopy every 5 years; S = sigmoidoscopy every 5 years; R = proto-type radiology procedure every 5 years; C = colonoscopy every 10 years; F = annual fecal occult blood test; WD = strategy is weakly dominated by at least one other strategy; SD = strategy is strongly dominated by at least one other strategy.
SOURCE: M. Pignone Workshop Presentation (Appendix I).

Comparing strategies with one another. Under standardized assumptions, all modules agreed about which strategies survived the dominance test (Table 7), but the incremental cost per year of life gained for F/S, versus S, still varied widely.

Effect of Specific Assumption Groups on Variations across Models

The research teams examined the separate effect of each group of assumptions on the estimates for four strategies (versus no screening).[9] They compared model results when each of the four assumption groups was standardized while the rest were set to their original values. Estimated years of life gained did not show any general pattern of convergence (Table 8). The ratio between the highest and lowest estimate of years of life gained actually increased for some strategies when some assumption groups were standardized. The range of estimates for lifetime costs associated with a particular strategy declined substantially for two strategies but increased slightly for two others.[10]

TABLE 8. Effect of Standardizing Individual Assumption Groups on Variation Across Models: Ratio of Highest Estimate to Lowest Estimate

	F/S	S	C	F
Net Present Value of Years of Life Gained Due to Screening				
Original assumptions	1.3	1.6	2.4	1.7
Costs std	1.3	1.6	2.4	1.7
Test performance std	1.2	1.8	1.4	1.4
FU/Surveillance std	1.3	1.7	1.5	1.7
Compliance std	1.4	1.8	1.3	1.3
All groups std	1.4	2.0	1.4	1.5
Net Present Value of Lifetime Costs Incurred Due to Screening				
Original assumptions	3.9	5.2	3.6	2.9
Costs std	2.8	3.2	4.1	NA
Test performance std	4.7	7.8	3.8	6.0
FU/Surveillance std	4.0	5.0	4.2	2.5
Compliance std	3.0	3.9	2.5	5.6
All standard	1.9	2.0	1.5	2.0

NOTES: Net present value computed at 3 percent per annum; FU/Surveillance = Follow up and Surveillance; std = standardized; NA = Not available. One model predicted net cost savings, thereby making calculation of a ratio impossible.

---

[9] Recall that strategy R was conceived with standard test performance and test costs at the outset because at the time not all of the research teams had studied a radiologic technology for screening. Therefore, the analysis of the impact of individual strategies did not include R.

[10] One model estimated negative net lifetime costs for strategy F under standardized cost assumptions. The absolute difference in costs between the highest-cost and lowest-cost estimates increased compared with the difference among the models when original assumptions were used.

TABLE 9. Effect of Standardizing Specific Assumption Groups on Variation in Cost-Effectiveness Ratios across Models: Ratio of Highest Estimate to Lowest Estimate

	Strategy				
	F/S	S	R	C	F
Original assumptions	3.7	3.3	5.2	4.0	2.4
Costs standard	3.1	2.3	5.6	2.5	NA
Test Performancestandard	4.4	6.2	5.5	3.7	5.9
FU/Surveillance standard	3.8	3.6	1.8	2.4	2.4
Compliance standard	3.5	3.8	5.7	3.2	5.3
All groups standard	1.8	2.1	1.8	1.7	1.9

NOTES: Cost-effectiveness compared with no screening. NA = Not available. One model predicted net cost savings, thereby making calculation of a ratio impossible.

The cost-effectiveness ratios for each strategy did converge across models as a result of standardizing costs (Table 9). That result led Michael Pignone to conclude that standardizing cost assumptions seemed to have the biggest effect on convergence among models. However, he also warned that standardizing other groups of assumptions individually did not lead to systematic convergence across models in the estimated cost-effectiveness of any strategy. Because the cost-effectiveness ratios converged when all four assumption groups were standardized, Pignone observed, it is probable that the assumption groups interact in their effects on model outcomes.

Lessons Learned from the Exercise

The results of the pre-workshop exercise prompted substantial discussion among the workshop participants. Comments focused both on the strengths and limitations of the exercise itself and on the implications of the collaborative exercise for further model development.

The Impact of Subtle Differences in Model Structure

Workshop participants identified some subtle differences in structure across models that affected the results of the exercise itself. One is how the different models account for polyps that are not adenomas. As described by T.R. Levin, most experts believe that the vast majority of colorectal cancers arise from pre-cancerous adenomas. These lesions come in a variety of morphologic and histological forms and they grow and progress to cancer at varying speeds. They are not, however, the only polyps that appear in the colon or rectum—other kinds of benign lesions, notably hyperplastic polyps, are quite common in older people (Lieberman et al., 2003). Although hyperplastic lesions are thought to present a low risk for progression to cancer (Imperiale et al., 2003; Lieberman et al., 2004), some screening technologies may detect them with higher frequency than others. In particular, endoscopy and radiology would be more likely to detect non-adenomatous polyps than would fecal occult blood testing, because non-adenomatous polyps rarely bleed.[11] Once detected, however, such lesions are typically removed and sent for biopsy because they cannot be differentiated from adenomas by any other method. Martin Brown observed that the cost of follow-up procedures triggered by detection of a non-adenomatous lesion may have a major effect on the incremental cost of screening.

---

[11] Even when a non-adenomatous polyp is not detectable by the screening test, it could be found serendipitously through diagnostic follow up of a test result that was positive for reasons unrelated to its presence.

TABLE 10. Impact of Excluding Non-Adenomatous Polyps from the Vanderbilt Model: Percent Change in Outcomes Resulting from Exclusion

Strategy	Change in Lifetime Cost Percent	Change in Years of Life Gained Percent
F/S	- 34	- 1
S	- 45	- 9
R	- 38	- 1
C	- 14	- 0
F	-14	- 0

NOTES: F/S = annual fecal occult blood test, sigmoidoscopy every 5 years; S = sigmoidoscopy every 5 years; R = prototype radiology procedure every 5 years; C = colonoscopy every 10 years; F = annual fecal occult blood test; WD = strategy is weakly dominated by at least one other strategy; and SD = strategy is strongly dominated by at least one other strategy.
SOURCE: R. Ness Workshop Presentation (Appendix G).

The research teams reported that not all of the models account for the implications of detecting non-adenomatous lesions. Karen Kuntz noted that the Harvard model did not include such lesions at all. Some recognize them implicitly rather than explicitly by making a downward adjustment in the assumed specificity (i.e., increasing the false positive rate) of the screening test, or an upward adjustment in the average cost of diagnostic follow-up of adenomas detected through screening. Reid Ness observed that standardizing assumptions in the groups involving test performance (i.e., test specificity) and costs (i.e., follow-up costs) masked these subtle differences in model structure.

The Vanderbilt team was the first to recognize the impact of non-adenomas on the standardized results of the pre-workshop exercise. Vanderbilt's estimates of the lifetime costs of all screening strategies were much higher than those reported to the workshop by the other research teams (see Figure 3B). The Vanderbilt model explicitly recognizes the prevalence of non-adenomatous polyps and independently records the costs of diagnostic follow-up of those lesions. Because other models either excluded those costs or considered them implicitly through adjustments in other assumptions, they effectively ignored them when test specificity and unit costs were standardized. The Vanderbilt team assessed the importance of this difference in model structure by reanalyzing the five strategies under fully standardized assumptions after setting the prevalence of non-adenomas to zero in their model. Reid Ness reported that the lifetime costs of all screening strategies declined (Table 10). Those with the highest relative decline were the screening strategies most likely to detect non-adenomas, namely those that involve direct visualization of the colon and diagnostic follow-up of all polyps with colonoscopy.[12] Ignoring non-adenomas also had a small negative impact on life years gained, because doing so would imply fewer referrals to colonoscopy. Such referrals generated by a screening test that was positive because of a non-adenomatous polyp would sometimes result in serendipitous discovery on follow-up of an adenoma or cancer, with consequent life-extending benefits.

---

[12] Note that under the standardized assumptions\ all polyps found on sigmoidoscopy are referred to full colonoscopy for removal and biopsy.

TABLE 11. Impact of Detecting Non-Adenomas on Incremental Cost-Effectiveness Ratios Study

Strategy	Harvard	Ladabaum	Miscan	Vijan	Vanderbilt (old)	Vanderbilt (new)
F/S	$99,977	$79,920	$55,878	$56,969	$355,647	$355,608
S	SD	SD	SD	SD	SD	SD
R	WD	SD	WD	SD	$209,906	$114,510
C	SD	SD	SD	SD	WD	SD
F	$11,632	$7,272	$5,980	$9,676	$10,073	$8,659

NOTES: F/S = annual fecal occult blood test, sigmoidoscopy every 5 years; S = sigmoidoscopy every 5 years; R = prototype radiology procedure every 5 years; C = colonoscopy every 10 years; F = annual fecal occult blood test; WD = strategy is weakly dominated by at least one other strategy; and SD = strategy is strongly dominated by at least one other strategy.
SOURCE: R. Ness Workshop Presentation (Appendix G).

The Vanderbilt team reported that their reanalysis had a limited effect on incremental cost-effectiveness ratios under standardized assumptions. Their results were in closer agreement with those of the other models, but their estimate of the incremental cost-effectiveness of moving from F to F/S was still much higher (Table 11). Ness concluded that different approaches to non-adenomas may have been responsible for some of the variation among models. However, other factors recognized but not fully understood by participants continued to support a high level of variation in the incremental cost-effectiveness estimates for alternative screening strategies.

Other Limitations of the Exercise

Several participants noted that standardizing to a single set of values in each assumption group is insufficient if the goal is to determine unequivocally the extent to which variation across models can be explained by different values in the four groups of assumptions.[13] Other standardized values for the same group of assumptions might have generated more, or less, agreement among models than did the values chosen for the pre-workshop exercise. In the extreme, it might be possible to force a measure of agreement among models by selecting standardized assumptions that strongly favor certain strategies. Judith Wagner noted that the standardized assumptions selected in the pre-workshop exercise may have differentially favored the two strategies involving fecal occult blood testing—F and F/S—since the effectiveness of fecal occult blood testing is especially sensitive to assumptions about compliance. A more robust exercise would have tested multiple values for standardized assumptions, perhaps selected probabilistically from a range of possible values. Such an exercise—involving hundreds or thousands of model runs—would have required time and resources that none of the research teams could afford without external funding.

Several participants noted that convergence among models does not necessarily imply that the models are valid representations of the true cost and effectiveness of any given CRC screening strategy. To paraphrase Marjolein van Ballegooigen, if the models merge when we standardize, should we believe the merged results? The ultimate test of any model is how well it predicts what occurs in the real world. If all models share flawed designs or assumptions, agreement does not constitute validity.

---

[13] Recall that natural history assumptions were not standardized in the exercise.

Michael Pignone observed that a key structural aspect shared by all the models is that the sensitivity of each subsequent screening test performed in a periodic screening program is independent of the results of earlier tests. That assumption may be questioned most strongly in the case of annual fecal occult blood testing. Researchers have posited that some adenomas may never bleed, while others may bleed regularly. If more were known about whether such patterns actually exist, and the frequency with which they do, models could be constructed that would adjust the assumed probability that people with adenomas receive positive fecal occult blood testing results in the second and subsequent years of a screening program, based on their test results in previous years. In Pignone's view, adjustments such as these could have profound effects on the estimated effectiveness of periodic fecal occult blood testing. At present, however, data simply do not exist to provide reasonable estimates of such contingent probabilities, and modeling them would be a complex undertaking.

MAJOR CHALLENGES TO MODELING THE COST-EFFECTIVENESS OF CRC SCREENING

The workshop benefited from presentations by leading researchers on the current state of knowledge about the natural history of CRC and the effects of screening, follow-up and surveillance. Those presentations took place on the afternoon of the first day and are published in the appendixes to this workshop report. They covered the following topics:

- Evidence on test performance of current and experimental screening technologies—Brian Mulhall, M.D.;
- Issues in measuring the costs of CRC screening and treatment—Martin Brown, Ph.D.;
- Evidence on rates of compliance with CRC screening—Sally Vernon, Ph.D.;
- Evidence on endoscopy utilization and capacity—Laura Seefe, M.D.;
- Evidence on compliance with follow up and surveillance—Todd Anderson, M.S.;
- Evidence on efficacy of follow-up and surveillance protocols—Deborah Schrag, M.D.;
- Evidence on the natural history of CRC—T.R. Levin, M.D.; and
- The National Cancer Institute's CISNET Program and its approach to model validation—Eric (Rocky) Feuer, Ph.D.

The presentations were intended to identify the best evidence both to improve models and to identify gaps in knowledge. Together with the collaborative modeling exercise, they stimulated workshop participants to confront three major issues that challenge the ability of models to provide useful information to health policy makers.

Uncertainty

Workshop participants spent a good deal of time addressing the uncertainty underlying the costs and effects of colorectal cancer screening. Louise Russell and Michael Pignone argued that pervasive uncertainty makes bottom-line conclusions about the comparative performance of different screening strategies, although not about the overall cost-effectiveness of screening itself, potentially inappropriate because they presume a degree of precision that the current state of knowledge cannot support and may never be able to support. Ironically, it is exactly that kind of precise, bottom-line guidance that decision makers seek.

The Extent of the Problem

The workshop presentations underscored how little is known about many aspects of screening or its consequences. Brian Mulhall's review of the wide range of estimates of fecal occult blood test sensitivity and specificity for adenomas in people who are recommended for screening suggested that uncertainty about test performance is not limited to new, emerging, or uncommon technologies. Fecal occult blood testing, one of the oldest technologies available for CRC screening, has been the focus of several large-scale randomized screening trials, all of which have demonstrated that it can reduce mortality from CRC (Jorgensen et al., 2002; Mandel et al., 1999; Scholefield et al., 2002). However, according to Mulhall, none of those trials has provided definitive evidence on its sensitivity and specificity for adenomas.

The range of uncertainty about the performance of other common screening technologies, such as sigmoidoscopy and colonoscopy, may be lower, but there is still substantial variation in findings across studies, according to Mulhall. Long considered the gold standard for detection of adenomas, with a sensitivity that was believed to be close to 100 percent, colonoscopy was found in a recent head-to-head comparison with virtual colonoscopy, a new radiological procedure, to miss about 11 percent of advanced adenomas (Pickhardt et al., 2003).

Martin Brown emphasized the uncertainty surrounding estimates of CRC treatment cost, which is a major component of the lifetime cost of a screening program. Variation in estimates of this cost, and its distribution over the years following diagnosis, can make for very large differences among models in estimates of the net cost of screening. A strategy that prevents a large number of cancers is far less costly when treatment costs, and thus savings from early detection, are high rather than when they are low. Although it might seem easy to make accurate estimates of such costs because billing and claims data are available from health care providers or insurers, prices can vary widely from provider to provider and across different payers. Moreover, estimates vary depending on whether they are based on prices charged or on audits of the amount and value of the labor and other inputs required to produce each service. Thus, Brown concluded, a seemingly straightforward element—the cost of treating colorectal cancer—is in practice subject to considerable uncertainty.

In her discussion of the current evidence on compliance with CRC screening, Sally Vernon emphasized the uncertainty about compliance in a screening program that continues over a patient's lifetime. Though much is known about factors that affect patients' adherence to screening, notably insurance coverage, education, and physician recommendations, survey evidence is insufficient to provide accurate estimates of the levels of adherence to periodic screening that could be expected over the long term. William Lawrence observed that we do not know whether compliance rates estimated from one-time

consumer surveys represent a combination of some patients who fully adhere to a lifelong screening program and others who never receive any screening, or whether they represent a more homogenous population all of whom adhere to a screening strategy intermittently. In Lawrence's view, these distinctions have important implications for the effectiveness and cost of screening programs.

Uncertainty about compliance is also high because surveys define compliance differently. Sally Vernon observed that surveys that measure the number of patients who receive fecal occult blood test kits from their physicians typically report high compliance, whereas those that measure the number of test kits returned for analysis show much lower rates.

Uncertainty about the natural history of colorectal cancer in average-risk individuals was also a topic of discussion throughout the workshop. T.R. Levin summarized the evidence on several important aspects of that history. His review of one aspect—the proportion of CRCs that arise *de novo*, without spending time as a pre-cancerous adenoma— illustrates how difficult it can be to resolve uncertainty. Cancers arising from fast-progressing adenomas or from adenomas that are difficult to detect with even the most sensitive screening tests could be mislabeled as *de novo*. The difference could be important for comparing strategies because, for example, underestimating the proportion of cancers that arise *de novo* could favor screening technologies that have high sensitivity for adenomas over those that are better at diagnosing early cancer. Brian Mulhall observed that the emergence of molecular assays in the near future may offer new opportunities for definitive research on the question of *de novo* cancer.

How to Think About Uncertainty

A prerequisite to dealing with the effects of uncertainty is to recognize that it comes in different forms. Rocky Feuer offered a three-level classification. The first type, which he referred to as stochastic variation, arises from inherent randomness in disease processes and human behavior across the members of a population. Put simply, not everyone's disease follows the same course and screening and treatment do not have the same effects from person to person. Dealing with stochastic variation by itself is straightforward. Modelers would simply specify the known distribution of values for an input parameter. Statistical confidence intervals for model outcomes can be generated through analytic or simulation techniques. Feuer pointed out that the uncertainty discussed by workshop speakers and participants does not fall into this category.

The second level—parameter uncertainty—refers to the far more common situation in which the true values of the parameter, e.g., test sensitivity or the cost of treating CRC, are not well understood. Estimates about the population distributions of model inputs, drawn from medical and epidemiological research, are the "assumptions" on which models are built. Feuer explained that sensitivity analysis is the most appropriate approach to dealing with this kind of uncertainty. In sensitivity analysis, modelers let assumptions vary across a range of likely values and the resulting range of costs and effects is reported. When many parameters are uncertain, as they are in the case of colorectal cancer screening, experts recommend the use of probabilistic cost-effectiveness analysis to generate a bottom-line "confidence interval" or "credible interval" of cost-effectiveness ratios (Briggs et al., 2002; Gold et al., 1996). Such an interval allows users to understand the simultaneous effect of uncertainty about many parameters on the range of cost-effectiveness ratios that result.

The third level of uncertainty is structural, according to Feuer. In that case, researchers may have little to go on about the relationships and interactions among key parameters. They may therefore choose to model those relationships in different, perhaps even arbitrary ways. The debate over whether some cancers arise *de novo* or from pre-existing adenomas is an example of structural uncertainty. To deal with this unknown, some modelers have assumed that such lesions are simply very fast-moving adenomas, while others have assumed that they can never be detected as adenomas. Other examples of structural uncertainty are the effects of including or excluding the consequences of detecting non-adenomatous polyps, the cost of a patient's time engaged in screening, or the impacts on individuals' quality of life from both screening and colorectal cancer. In Feuer's view, the five models highlighted in this workshop represent five different approaches to resolving structural uncertainty.

Strategies for Managing Uncertainty

As workshop participants grappled with how best to deal with the effects of parameter and structural uncertainty on the ability of CEA models to produce the answers policy makers want, many ideas surfaced on how to reduce, or at least manage, those effects.

Research strategies. To reduce the uncertainty about important assumptions, many participants called for more primary research, particularly on those factors that account for the greatest variation among models. The areas most frequently cited were costs and compliance. Laura Seefe outlined work that the CDC is conducting with several states to enhance the utilization of CRC screening. That research should provide more evidence on the degree of compliance that can be expected from different screening program designs. Alan Gerling endorsed more studies of the impact of public awareness programs and other recruitment strategies on adherence, along the lines of those currently underway in pilot studies in the United Kingdom. Michael Pignone suggested that research aimed at getting better estimates of the lifetime cost of treating colorectal cancer might do more to resolve differences among models than would research on other parameters.

Another approach mentioned by several workshop participants to help understand the effect of uncertainty is to evaluate model predictions against independent results from well-designed trials of screening programs. The presentations by leaders of the five research teams showed that several have used data from large fecal occult blood testing screening trials to evaluate the extent to which their models' predictions of cancer incidence and mortality over time agree with the results found in the trials. The ongoing PLCO trial (Schoen et al., 2003; Gohagan et al., 1995), which is testing sigmoidoscopy screening, will soon provide a new dataset to support validation, according to Robert Schoen. NCI's CISNET program acts as a catalyst for sharing useful databases from NCI-sponsored studies among member research teams, said Rocky Feuer.

Karen Kuntz reminded the workshop participants that the paucity of data on important assumptions often leads model builders to use data from trials to inform their choice of values for critical assumptions, such as the sensitivity and specificity of screening tests. She warned that validating a model with data that were also used in part to build model assumptions does not provide a true test of the validity of model predictions.

Short of evaluating the predictive validity of models with independent data sources, research teams can assess other measures of validity,[14] such as whether a model contains all of the components of cost and effect that one would expect to be important. For ex-

---

[14] For a description of different kinds of validity, see the research methods web page maintained by William Trochim at http://www.socialresearchmethods.net/kb/index.htm (Trochim, 2004).

ample, Reid Ness's presentation of his team's reanalysis of the pre-workshop exercise revealed that leaving out the cost of working up non-adenomatous polyps can have a large effect on model outcomes. Louise Russell mentioned that an important component of cost omitted in all of the CRC models presented at the workshop is the value of patients' time lost in screening. That omission biases model results toward screening technologies that are time intensive for the patient and against technologies that are fast and convenient. Martin Brown explained that with no published empirical studies of this cost component modelers typically exclude it.

Several participants expressed skepticism that either research approach—primary research on uncertain assumptions or excluded components or greater availability of independent data sets for model validation—will fully resolve uncertainty, in part because of the cost of generating new information but also because technological advances in screening and treatment continually create new unknowns. In technical terms, the confidence or credible interval for the cost-effectiveness of one strategy is likely to overlap that of others. Recognizing this reality, they suggested steps that might help decision makers make more appropriate use of the information that CEAs can generate.

User strategies. One line of thinking expressed at the workshop was that policy makers have little choice but to accept the discrepant results from models because those results simply reflect the lack of medical and epidemiological evidence. Policy makers could adopt a message that focuses on the value of CRC screening in general, leaving specific choices of strategies to physicians and patients. Robert Dittus suggested that an appropriate message for providers might be, "here is a collection of approaches that we think are good and they all fit within the general realm of 'it's a whole lot better than nothing." Robert Smith, on the other hand, argued for opinion leaders to advocate a practical screening strategy that offers the greatest protection and best outcomes for patients given what we know today.

Others noted that colorectal cancer screening involves high cost as well as great medical benefits, so choosing one strategy over another can mean differences of billions of dollars and hundreds of thousands of years of life when summed over the entire U.S. population 50 years of age and older. Although Richard Lilford observed that "the complexity of choice is so great in medicine that the questions are unanswerable," he also recognized that decisions must nevertheless be made based on the best information available. In his view, models offer one type of information to assist in those decisions. The ultimate choice of screening strategy, according to Lilford, will be influenced by political pressures and preferences as well as by models laying out costs and effects as best they can.

Some participants addressed ways to help policy makers better understand the levels of uncertainty represented in models. Judith Wagner commented that editors of medical journals can play a useful role in this regard. The pre-workshop exercise revealed, for example, how uneven and sometimes vague the descriptions in published papers were of models' assumptions about compliance and diagnostic follow-up protocols. Clear descriptions of the assumptions in these and other important areas should be a priority. Wagner also observed that published CEAs of CRC screening have often evaluated a single screening strategy not examined in published work by other modeling teams. That practice makes it difficult for readers to assess the level of agreement across models. Authors seeking to evaluate the cost-effectiveness of a new screening technology might be asked to report on the model's outcomes for a common set of well studied screening strategies, such as the five strategies used in the pre-workshop exercise. Decision makers

could then assess whether the cost-effectiveness results for the new technology might differ if assessed by other models.

Some participants called for research teams to provide more access to their models, even suggesting that models be placed in the public domain on the Internet to allow decision makers to test the impact of different assumptions or strategies. Others pointed out, however, that models entail a substantial investment in researchers' intellectual capital, which could be compromised by open access. Rocky Feuer suggested as a middle ground the Model Profiler currently under development as part of the CISNET program. The Profiler, described in Feuer's presentation, is expected to provide open web-based access to detailed information on model structure, assumptions and outputs, but not to the models themselves.

Modeling Reality or an Ideal World?

Another issue that threaded its way through the discussion concerned which of the following two questions CRC screening models should seek to answer: "What can be expected to happen under a given strategy?" Or "What could happen if the strategy is implemented under ideal conditions?"

The modeling of compliance is an obvious example of this issue. Perfect adherence to a strategy, including the periodic screening examinations, the specified follow-up, and surveillance protocols is the ideal, but it is not achieved in practice, as Sally Vernon showed in her presentation.

Another example of the tension between modeling the ideal and modeling reality is how models handle diagnostic follow-up of a positive sigmoidoscopy. Some experts recommend that polyps found on sigmoidoscopy be biopsied or removed during that procedure, with referral for a full colonoscopy only if the polyp is found to be a high-risk or advanced adenoma. But Todd Anderson's presentation on the frequency of follow-up and surveillance procedures suggests that the vast majority of polyps removed from patients undergoing sigmoidoscopy are removed in a subsequent colonoscopy. Whether models are based on recommended or actual practice in this regard could affect the costs and effects of sigmoidoscopy.

Several participants noted that limits on the supply of screening procedures represent an area in which reality may force departures from ideal conditions. Certain procedures—notably colonoscopy, but in the future, perhaps, virtual colonoscopy—require trained specialists to perform or interpret them. Like most people, medical specialists respond to economic incentives and higher reimbursements for screening or surveillance procedures would induce them to do more. However, some strategies may require so many colonoscopies that the supply of gastroenterologists would be completely inadequate.[15] The same may be true of radiologists, should virtual colonoscopy become a routine screening procedure (Herdman and Norton, 2005). The supply of specialists cannot be expanded quickly, so real constraints on capacity may have to be taken into account. Sandeep Vijan observed that assuming low compliance for certain screening or surveillance procedures is one way models could implicitly account for such constraints.

Seth Glick emphasized the divergence between test performance under ideal quality assurance programs and test performance in current practice. The quality of many screening processes may be poor today, in Glick's view. Estimates of test sensitivity, specificity, and medical risk, usually taken from studies where good quality assurance existed,

---

[15] Recent news accounts suggest that waiting times for colonoscopy are growing in certain areas of the country (Kowalczyk, 2004). New evidence also suggests that some surveillance colonoscopies are being done more frequently than is recommended by professional guidelines (Mysliwiec et al., 2004).

therefore overestimate the performance of screening in the absence of strong quality assurance programs.

The basic problem, in the view of Martin Brown, is not the choice between the real and the ideal per se, but the failure of researchers to make explicit their choices about it. Moreover, their implicit choices may vary within a model, with optimal assumptions in one area and realistic assumptions in another and no explicit rationale for the differences. He and several other participants argued that both kinds of analysis are useful. For example, analysts could tell patients and physicians what can be expected if consumers fully comply with a strategy, and then show the decrements in effects and costs resulting from less complete compliance. Michael Pignone cautioned that researchers who model ideal conditions need to include the full costs of achieving those conditions. Quality assurance and high rates of compliance do not come free. They are usually the result of intensive programs of behavior change that must continue for the duration of a screening program.

How Complex Should Models Be?

Participants returned repeatedly throughout the workshop to the question of whether models should be capable of evaluating complex screening strategies. This question surfaced often because the workshop participants, including the modelers, are fundamentally interested in the health policy question—what screening strategy is best?—not in modeling for its own sake. When the high lifetime costs of some very effective screening strategies become apparent in all models, a natural next step is to explore how those costs could be reduced, without compromising effectiveness, by fine-tuning strategies. Such fine-tuning drives modelers to add more branches to their strategies, which places even greater demands on the clinical and epidemiological evidence available to support such modeling.

Many ideas for complex screening strategies were put forward at the workshop. Participants suggested strategies such as changing the screening, follow-up, or surveillance schedule as a person ages, offering different screening strategies to different demographic groups with different relative risks, or changing a schedule contingent on the results of a previous screening, follow-up, or surveillance test. For example, Ann Zauber observed that men and women have different profiles of adenoma and CRC incidence, with women developing CRC an average of 10 years later than men do (Chu et al., 1994; Cooper et al., 1995; Devesa and Chow, 1993). Different screening strategies for men and women might make sense and could be explored by models. Donatus Ekwueme raised similar possibilities for tailoring strategies by race in recognition of the systematic racial differences in incidence, prevalence, and location in the colon of adenomas and polyps (Devesa and Chow, 1993; Theuer et al., 2001; Walker Jr et al., 1995). Michael Pignone and Marjolein van Ballegooijen suggested that it might make sense to alter the type of screening test as a person ages, saving more sensitive but more expensive tests until the individual is at higher risk of advanced adenomas or CRC.

To John Inadomi, the most important clinical question is whether surveillance following polypectomy is cost-effective and how often it is needed. He and Reid Ness argued that selective post-polypectomy surveillance strategies—where high-risk individuals are monitored more often than those at low-risk of future polyps or cancer—have the greatest potential to reduce costs. But, to make such judgments without imperiling outcomes, accurate data on the factors that matter are needed. Deborah Schrag summarized the results of the National Polyp Study (Winawer et al., 1993), which found that intensive surveillance strategies offer little additional benefit compared with protocols that condi-

tion future surveillance on the nature of the polyp removed and on the result of the first surveillance test.

Mark Fendrick raised the possibility of adjusting screening strategies to account for individuals who had already received a colonoscopy for symptoms or non-screening reasons. David Lieberman commented that in reviewing a large endoscopic database he found 40 to 50 percent of all colonoscopies were performed either for vague symptoms or for rectal bleeding. The results of most of those procedures either are negative or show benign polyps. If models were adjusted to assume that 40 to 50 percent of individuals have already been screened before a formal screening program starts, rather than assuming that no one has been screened, as they currently do, the predicted cost of screening would be lower.

Despite the enthusiasm for complex strategies (and for estimating their potential to save costs or increase effectiveness), several participants sounded notes of caution. Mark Fendrick emphasized barriers to making complex screening strategies operational. He described the difficulty one major medical center had in providing same-day follow-up colonoscopy for people with positive screening sigmoidoscopic examinations, even when those patients had been clinically prepped beforehand for a possible colonoscopy. He and Sandeep Vijan also warned that presenting too many options could overwhelm patients and ultimately reduce their willingness to participate in screening or surveillance. Robert Dittus held out hope that new medical information systems, such as automated test ordering and electronic medical records with built-in guidelines, will make it easier for physicians to implement complex strategies.

Several participants held that if complex strategies offer substantial hope for moderating costs without reducing the benefits of screening, then models should stand ready and be capable of assessing them. But, argued Amnon Sonnenberg, given the information requirements of complex models, we may be expecting models to do too much. Sometimes if models become too complex, they go off the mark simply because they must make too many assumptions based on too little evidence. Thus, the discussion of complexity ended with a reprise of the first problem for modeling, as for decision making in general: uncertainty.

NEXT STEPS

The workshop was not intended to evaluate the effectiveness or cost-effectiveness of colorectal cancer screening in general. Virtually all economic models, drawing from a wealth of clinical trials and epidemiological studies, have found that colorectal cancer screening decreases mortality from the disease. The message to physicians, payers, and patients that periodic screening for colorectal cancer is an effective preventive measure continues to have urgency.

The focus of the workshop was not on "whether to screen" but on "how to screen." Nevertheless, it was NOT intended to evaluate alternative CRC screening strategies. No evidence was presented to recommend one strategy over others. Instead, the purpose was to explore and enhance the usefulness of cost-effectiveness models in helping medical policy makers make such judgments. An obvious next step would be for modelers and clinicians to continue to explore together how the factors affecting model outcomes that were identified in this workshop—both parameter assumptions and structural assumptions—can be resolved through better information or better modeling. Such explorations take time and resources beyond those available to a one-time workshop. The workshop did show, however, that collaboration can identify critical sources of variation—such as

assumptions about the cost of treating disease—that lead to conflicting findings. Those differences confuse decision makers, who must grapple with the underlying question of "which is the best strategy?"

Models are like maps. Maps are useful when they serve as guides to underlying territories. A map that is too vague is useless; one that is completely accurate merges with the territory itself and is also useless as a guide. The participants spent considerable time discussing the optimal balance for models along the continuum from rough guide to complete accuracy. They struggled with questions of how detailed CRC models should be if they are to be useful to decision makers and how detailed they can be, given the available information.

Richard Lilford provided valuable perspective with his observation that "modeling is a way of having a conversation." That is precisely what occurred during the day and a half when modeling teams and experts came together to compare assumptions, results, and the underlying evidence base for modeling. Many participants commented on the value of the conversation for further refinement of their models (in the case of the research teams) and for research ideas (in the case of clinical and epidemiological researchers). The pre-workshop modeling collaboration demonstrated that too many lives and dollars are at stake not to continue to work on understanding and communicating both the strengths and weaknesses of cost-effectiveness models.

REFERENCES

Briggs AH, O'Brien BJ, Blackhouse G. 2002. Thinking outside the box: Recent advances in the analysis and presentation of uncertainty in cost-effectiveness studies. *Annu Rev Public Health* 23: 377-401.

Chu KC, Tarone RE, Chow WH, Hankey BF, Ries L. 1994. Temporal patterns in colorectal cancer incidence, survival, and mortality from 1950 through 1990. *J Natl Cancer Inst* 86(13): 997-1006.

Cooper GS, Yuan Z, Landefeld CS, Johanson JF, Rimm AA. 1995. A national population-based study of incidence of colorectal cancer and age: Implications for screening in older Americans. *Cancer.* 75(3): 775-781.

Cotton PB, Durkalski VL, Pineau BC, Palesch YY, Mauldin PD, Hoffman B, Vining DJ, Small WC, Affronti J, Rex D, Kopecky KK, Ackerman S, Burdick JS, Brewington C, Turner MA, Zfass A, Wright AR, Iyer RB, Lynch P, Sivak MV, Butler H. 2004. Computed tomographic colonography (virtual colonoscopy): A multicenter comparison with standard colonoscopy for detection of colorectal neoplasia. *JAMA* 291(14): 1713-1719.

Curry S, Byers T, Hewitt M (eds). 2003. *Fulfilling the Potential of Cancer Prevention and Early Detection.* National Cancer Policy Board Institute of Medicine. Washington DC: The National Acadamies Press.

Devesa SS, Chow WH. 1993. Variation in colorectal cancer incidence in the United States by subsite of origin. *Cancer* 71(12): 3819-3826.

Eddy DM. 1980. *Screening for Cancer: Theory, Analysis and Design.* Englewood Cliffs, NJ: Prentice-Hall.

Eddy DM, Nugent FW, Eddy JF, Coller J, Gilbertsen V, Gottlieb LS, Rice R, Sherlock P, Winawer S. 1987. Screening for colorectal cancer in a high-risk population. Results of a mathematical model. *Gastroenterol* 92(3): 682-692.

Edwards BK, Howe HL, Ries LAG, Thun MJ, Rosenberg HM, Yancik R, Wingo PA, Jemal A, Feigal EG. 2002. Annual report to the nation on the status of cancer, 1973-1999, featuring implications of age and aging on U.S. cancer burden. *Cancer* 94(10): 2766-2792.

Frazier AL, Colditz GA, Fuchs CS, Kuntz KM. 2000. Cost-effectiveness of screening for colorectal cancer in the general population. *JAMA* 84(15): 1954-1961.

Glick S, Wagner JL, Johnson CD. 1998. Cost-effectiveness of double-contrast barium enema in screening for colorectal cancer. *AJR AM JRoentgeno* 170(3): 629-636.

Gohagan JK, Prorok PC, Kramer BS, Hayes RB, Cornett JE. 1995. The prostate, lung, colorectal, and ovarian cancer screening trial of the National Cancer Institute. *Cancer* 75(7 Suppl.): 1869-1873.

Gold MR, Siegel J, Russell L, Weinstein, M. 1996. *Cost-Effectiveness in Health and Medicine* New York: Oxford University Press.

Herdman R, Norton L (eds). 2005. *Saving Women's Lives: Strategies for Improving Breast Cancer Detection and Diagnosis*. National Cancer Policy Board, Institute of Medicine. Washington DC: The National Acadamies Press.

Imperiale TF, Wagner DR, Lin CY, Larkin GN, Rogge JD, Ransohoff DF. 2003. Using risk for advanced proximal colonic neoplasia to tailor endoscopic screening for colorectal cancer. *Ann Intern Med* 139(12): 959-965.

Jorgensen OD, Kronborg O, Fenger C. 2002. A randomised study of screening for colorectal cancer using faecal occult blood testing: Results after 13 years and seven biennial screening rounds. *Gu.* 50(1): 29-32.

Joseph AM, Crowson TW, Rich EC. 1988. Cost effectiveness of HemoQuant versus Hemoccult for colorectal cancer screening. *J Gen Intern Med* 3(2): 132-138.

Khandker RK, Dulski JD, Kilpatrick JB, Ellis RP, Mitchell JB, Baine WB. 2000. A decision model and cost-effectiveness analysis of colorectal cancer screening and surveillance guidelines for average-risk adults. *Int J Technol Assess Health Care* 16(3): 799-810.

31

Kowalczyk L. 2004, August 1. Colonoscopies tax health industry. *The Boston Globe* A-1.

Ladabaum U, Song K, Fendrick AM. 2004. Colorectal neoplasia screening with virtual colono-scopy: When, at what cost, and with what national impact? *Clinical Gastroenterology & Hepatology* 2(7): 554-563.

Lieberman DA. 1995. Cost-effectiveness model for colon cancer screening. *Gastroenterol* 109(6): 1781-90.

Lieberman DA, Prindiville S, Weiss DG, Willett W. 2003. Risk factors for advanced colonic neo-plasia and hyperplastic polyps in asymptomatic individuals. *JAMA* 290(22): 2959-2967.

Lieberman DA, Prindiville S, Weiss DG, Willett W, Imperiale T. 2004. Risk factors for advanced colonic neoplasia: From evidence to application. *Evidence-Based Gastroenterology* 5(3): 86-87.

Loeve F, Boer R, Van Oortmarssen GJ, Van Ballegooijen M, Habbema JDF. 1999. The MISCAN-COLON simulation model for the evaluation of colorectal cancer screening. *Comput Biomed Res* 32(1): 13-33.

Loeve F, Brown ML, Boer R, Van Ballegooijen M, Van Oortmarssen GJ, Habbema JDF. 2000. Endoscopic colorectal cancer screening: A cost-saving analysis. *J Natl Cancer Inst* 92(7): 557-563.

Mandel JS, Church TR, Ederer F, Bond JH. 1999. Colorectal cancer mortality: Effectiveness of biennial screening for fecal occult blood. *J Natl Cancer Inst* 91(5): 434-437.

Mysliwiec PA, Brown ML, Klabunde CN, Ransohoff DF. 2004. Are physicians doing too much colonoscopy? A national survey of colorectal surveillance after polypectomy. *Ann Intern Med* 141(4).

Neilson AR, Whynes DK. 1995. Cost-effectiveness of screening for colorectal cancer: A simula-tion model. *IMA J Math Appl Med Biol* 12(3-4): 355-367.

Ness RM, Holmes AM, Klein R, Dittus R. 2000. Cost-utility of one-time colonoscopic screening for colorectal cancer at various ages. *Am J Gastroenterol* 95(7): 1800-1811.

Pickhardt PJ, Choi JR, Hwang I, Butler JA, Puckett ML, Hildebrandt HA, Wong RK, Nugent PA, Mysliwiec PA, Schindler WR. 2003. Computed tomographic virtual colonoscopy to screen for colorectal neoplasia in asymptomatic adults. *N Engl J Med* 349(23): 2191-2200.

Pignone M, Saha S, Hoerger T, Mandelblatt J. 2002. Cost-effectiveness analyses of colorectal cancer screening: A systematic review for the U.S. Preventive Services Task Force. *Ann Intern Med* 137(2): 96-104.

Ransohoff DF. 2004. Virtual colonoscopy: What it can do vs what it will do. *JAMA* 291(14): 1772-1774.

Schoen RE, Pinsky PF, Weissfeld JL, Bresalier RS, Church T, Prorok P, Gohagan JK. 2003. Re-sults of repeat sigmoidoscopy 3 years after a negative examination. *JAMA* 290(1): 41-48.

Scholefield JH, Moss S, Sufi F, Mangham CM, Hardcastle JD. 2002. Effect of faecal occult blood screening on mortality from colorectal cancer: Results from a randomised controlled trial. *Gut* 50(6): 840-844.

Smith RA, Cokkinides V, Eyre HJ. 2004. American Cancer Society guidelines for the early detec-tion of cancer, 2004. *CA Cancer J Clin* 54(1): 41-52.

Song K, Fendrick AM, Ladabaum U. 2004. Fecal DNA testing compared with conventional colo-rectal cancer screening methods: A decision analysis. *Gastroenterol* 126(5): 1270-1279.

Sonnenberg A, Delco F. 2002. Cost-effectiveness of a single colonoscopy in screening for colo-rectal cancer. *Arch Intern Med* 162(2): 163-168.

Sonnenberg A, Delco F, Inadomi JM. 2000. Cost-effectiveness of colonoscopy in screening for colorectal cancer. *Ann Intern Med* 133(8): 573-584.

Theuer CP, Wagner JL, Taylor TH, Brewster WR, Tran D, McLaren CE, Anton-Culver H. 2001. Racial and ethnic colorectal cancer patterns affect the cost-effectiveness of colorectal cancer screening in the United States. *Gastroenterol* 120(4): 848-856.

Trochim WM. 2004.The Research Methods Knowledge Base. 2nd ed. [online] http://www.socialresearchmethods.net/kb/index.htm.

United States General Accounting Office. 2004. Coverage of Common Colorectal Cancer Screen-ing Tests Is Common. Washington DC: United States General Accounting Office.

U.S. Preventive Services Task Force. 2002. Screening for colorectal cancer: Recommendation and rationale. *Ann Intern Med* 137(2): 132-141.

Vijan S, Hwang EW, Hofer TP, Hayward RA. 2001. Which colon cancer screening test? A comparison of costs, effectiveness, and compliance. *Am J Med* 111(8): 593-601.

Wagner JL, Herdman RC, Wadhwa S. 1991. Cost effectiveness of colorectal cancer screening in the elderly. *Ann Intern Med* 115(10): 807-817.

Wagner J, Tunis S, Brown M, Ching A, Almeida R. 1996. Cost-Effectiveness of Colorectal Cancer Screening in Average-Risk Adults. In: Young G, Rosen P, Levin B (eds). *Prevention and Early Detection of Colorectal Cancer.* Philadelphia: Saunders. Pp. 321-356.

Walker Jr B, Figgs LW, Zahm SH. 1995. Differences in cancer incidence, mortality, and survival between African Americans and whites. *Environ Health Perspect* 103(Suppl. 8): 275-281.

Whynes DK. 2004. Cost-effectiveness of screening for colorectal cancer: Evidence from the Nottingham faecal occult blood trial. *J Med Screen* 11(1): 11-15.

Winawer S, Fletcher R, Rex D, Bond J, Burt R, Ferrucci J, Ganiats T, Levin T, Woolf S, Johnson D, Kirk L, Litin S, Simmang C. 2003. Colorectal cancer screening and surveillance: Clinical guidelines and rationale: Update based on new evidence. *Gastroenterol* 124(2): 544-560.

Winawer SJ, Zauber AG, May Nah Ho, O'Brien MJ, Gottlieb LS, Sternberg SS, Waye JD, Schapiro M, Bond JH, Panish JF, Ackroyd F, Shike M, Kurtz RC, Hornsby-Lewis L, Gerdes H, Stewart ET, Lightdale CJ, Edelman M, Fleisher M. 1993. Prevention of colorectal cancer by colonoscopic polypectomy. *N Engl J Med* 329(27): 1977-1981.

Appendix A

Workshop Agenda:
Economic Models of Colorectal Cancer Screening in Average-Risk Adults

Lecture Room
National Academy of Sciences Building
2101 Constitution Avenue NW
Washington, D.C. 20418

January 26-27, 2004

Day 1
8:30 – 9:00   **Continental Breakfast,** Lecture Room

9:00 – 9:30   **Welcome and Introduction to the Workshop**
                Louise Russell, Ph.D., Rutgers University, Workshop Chair
                Martin Brown, Ph.D., National Cancer Institute, Workshop Co-Chair
                Judith Wagner, Ph.D., IOM Staff Officer

                PART I: COMPARING ECONOMIC MODELS OF CRC SCREENING

9:30 – 11:00   **Overview of Five Leading Economic Models of CRC Screening**
                The "Harvard Model": Karen Kuntz, Sc.D., Harvard University
                The "Ladabaum Model": Uri Ladabaum, M.D., University of California
                    at San Francisco
                The "Miscan Model": Marjolein van Ballegooijen, M.D., Erasmus University
                The "Vanderbilt Model": Reid Ness, M.D., Vanderbilt University
                The "Vijan Model": Sandeep Vijan, M.D., University of Michigan

11:00 – 11:15   **Break**

11:15 – 11:45   **Results of a Pre-Workshop Comparative Modeling Exercise**
                Michael Pignone, M.D., University of North Carolina

11:45 – 12:45   **Discussion**

12:45 – 1:15   **Lunch**

                **PART II: CURRENT EVIDENCE ON COLORECTAL
                CANCER SCREENING**

1:15 – 2:30   **Evidence on Compliance (Q&A follows presentations)**

                Current Evidence on Compliance
                Sally Vernon, Ph.D., University of Texas at Houston

35

Recent Follow-up and Compliance Patterns of Medicare Patients
Todd Anderson, M.S., Congressional Budget Office

The Relationship Between Compliance and Capacity
Laura Seeff, M.D., Centers for Disease Control

2:30 – 3:15   **Recent Findings on Test Performance**
Brian Mullhall, M.D., Walter Reed Army Hospital
Q&A follows presentation

3:15 – 3:30   **Break**

3:30 – 4:15   **Measuring the Cost of Screening, Follow-up, and Surveillance in Cost-Effectiveness Models**
Martin Brown, Ph.D. National Cancer Institute
Q&A follows presentation

4:15 – 5:00   **Evidence on Effectiveness of Follow-up and Surveillance Strategies**
Deborah Schrag, M.D., Memorial Sloan Kettering Cancer Center
Q&A follows presentation

Day 2
8:30 – 9:00   **Defining Issues: A Reprise of First Day's Findings**
Martin Brown, Michael Pignone, and Judith Wagner

PART III: MODELING THE NATURAL HISTORY OF COLORECTAL CANCER

9:00 – 9:45   **Recent Evidence on Natural History**
T.R. Levin, M.D., Kaiser Permanente Division of Research
Q&A follows presentation

9:45 – 10:00   **The NCI CISNET Project: An Overview of Goals, Methods, and Current Status**
Eric Feuer, Ph.D., National Cancer Institute

10:00 – 10:15 **Break**

10:15 – 11:15 **Discussion of Issues in Modeling Natural History**

11:15 – 12:00 **Modelers' Forum: Next Steps, Unresolved Questions, Critical Information Needs**
Marjolein van Ballegooijen, Karen Kuntz, Sandeep Vijan, Reid Ness, Uri Ladabaum and Coauthors

12:00 – 1:00   **Discussion and Wrap-up**

1:00   **Adjourn**

Appendix B

Workshop Participants

Todd Anderson, M.S.
Congressional Budget Office
Washington, DC

Elise Berliner, Ph.D.
Agency for Healthcare Research
  and Quality
Rockville, MD

Erica S. Breslau, Ph.D.
National Cancer Institute
Bethesda, MD

Durado Brooks, M.D.
American Cancer Society
Atlanta, GA

Martin L. Brown, Ph.D.
National Cancer Institute
Bethesda, MD

Kathleen Connors
"The Gray Sheet"
Chevy Chase, MD

Sarah Dash, M.P.H.
National Cancer Institute
Bethesda, MD

Robert S. Dittus, M.D.
Vanderbilt University Medical Center
Nashville, TN

Mary Doroshenk
American Cancer Society
Washington, DC

Donatus U. Ekwueme, Ph.D.
U.S. Centers for Disease Control
  and Prevention
Atlanta, GA

A. Mark Fendrick, M.D.
University of Michigan Health System
Ann Arbor, MI

Eric Feuer, Ph.D.
National Cancer Institute
Bethesda, MD

Alan Girling, M.D.
University of Birmingham
United Kingdom

Seth N. Glick, M.D.
Presbyterian Medical Center
Philadelphia, PA

Jay Herson
Westat
Rockville, MD

Inku Hwang, M.D.
Walter Reed Army Medical Center
Washington, DC

John M. Inadomi, M.D.
University of Michigan Medical Center
Ann Arbor, MI

Karen Kuntz, Sc.D.
School of Public Health
  Harvard University
Boston, MA

Uri Ladabaum, M.D., M.S.
University of California, San Francisco
San Francisco, CA

William Lawrence, M.D.
Agency for Healthcare Research
  and Quality
Rockville, MD

37

Theodore R. Levin, M.D.
Kaiser Permanente Medical Group
Oakland, CA

David Lieberman, M.D.
Oregon Health and Sciences
   University
Portland, OR

Richard Lilford, M.D.
University of Birmingham
United Kingdom

Diane Manninen, Ph.D
Battelle
Seattle, WA

Dion Morton, M.D.
University of Birmingham
United Kingdom
`
Brian P. Mulhall, M.D.
Walter Reed Army Medical Center
Washington, DC

Reid M. Ness, M.D.
Vanderbilt University
   Medical Center
Nashville, TN

Michael Pignone, M.D., M.P.H
University of North Carolina-
Chapel Hill
Chapel Hill, NC

George Provenzano, Ph.D.
Battelle
Arlington, VA

Louise B. Russell, Ph.D.
Rutgers University
New Brunswick, NJ

Howard Saft, M.D.
VA Medical Center
Washington, DC

Robert Schoen, M.D.
University of Pittsburgh
Pittsburgh, PA

Laura Seeff, M.D.
Centers for Disease Control
   and Prevention
Atlanta, GA

Robert A. Smith, Ph.D.
American Cancer Society
Atlanta, GA

Amnon Sonnenberg, M.D., M.Sc.
Portland VA Medical Center
Portland, OR

Deborah Schrag, M.D., M.P.H.
Memorial Sloan-Kettering Cancer
   Center
New York, NY

Marjolein van Ballegooijen, M.D.
Erasmus Universiteit
Rotterdam, The Netherlands

Sally W. Vernon, Ph.D.
University of Texas-Houston Health
   Science Center
Houston, TX

Sandeep Vijan, M.D.
University of Michigan
   Health System
Ann Arbor, MI

Iris Vogelaar, M.Sc.
Erasmus Medical Center
Rotterdam, The Netherlands

Judith L. Wagner, Ph.D.
Institute of Medicine
Washington, DC

Ann Zauber, Ph.D.
Memorial Sloan Kettering Cancer
   Center
New York, NY

Appendix C

Workshop Speaker and Staff Biographies

Louise B. Russell, Ph.D., (*Chair*) is Research Professor of Health Economics at the Institute for Health, Health Care Policy, and Aging Research, Rutgers University. Before coming to Rutgers in 1987, Dr. Russell was a Senior Fellow at the Brookings Institution for 12 years. She is an elected member of the Institute of Medicine of the National Academy of Sciences and currently serves on its National Cancer Policy Board. She co-chaired the U.S. Public Health Service Panel on Cost-Effectiveness in Health and Medicine, which published recommendations for improving the quality and comparability of cost-effectiveness studies in mid-1996 (*Cost-Effectiveness in Health and Medicine,* Oxford University Press; and a series of three articles in *The Journal of the American Medical Association,* October 1996). She was also a member of the first U.S. Preventive Services Task Force of the Department of Health and Human Services (1984-1988). She has published numerous articles and seven books, including *Educated Guesses: Making Policy About Medical Screening Tests* (1994), *Medicare's New Hospital Payment System: Is It Working?* (1989), and *Is Prevention Better than Cure?* (1986).

Martin L. Brown, Ph.D., (*Co-Chair*) has been the Chief of NCI's Health Services and Economics Branch since October 1999. Dr. Brown received his Ph.D. in Natural Resource Economics from the University of California, Berkeley, in 1981. He first joined the NCI in 1988 as a Health Economist, after serving as Associate Professor in the Department of Economics at Howard University. Since coming to NCI his research has focused on the economic burden of cancer to individuals and society; the acquisition and analysis of economic data on cancer prevention and control from controlled trials and from administrative records; evaluation and development of methodology and analytical tools for modeling and estimating the cost-effectiveness of specific cancer prevention and control interventions; programs and policies; the relationship of socioeconomic status, community structure, and health system organization to access to and use of cancer prevention and control services; and the financial structure of research support in the context of the changing system of healthcare delivery organization and financing.

Dr. Brown has published over 50 articles, reports, and book chapters in these research areas since coming to NCI. Dr. Brown is the economics editor of the *Journal of the National Cancer Institute*, has served as a reviewer for many journals, and has served on peer-review committees for the National Cancer Institute of Canada and the California Cancer Research Program. Dr. Brown serves as NCI Program Director for the Cancer Research Network, a national consortium of research organizations, affiliated with health maintenance organizations that conduct population-based research on cancer epidemiology, prevention and control.

Todd Anderson, M.S., joined the Congressional Budget Office in 1999 and has worked primarily on projects related to Medicare. He is the principal data analyst for CBO's work examining the circumstances of high-cost beneficiaries in Medicare and has also worked on a model to estimate the impact of competition-based reform proposals on Medicare. In addition, Mr. Anderson is the primary author of a CBO Study of pharmacy margins on Medicaid prescription drugs and has worked on cost estimates of proposals to provide wage insurance to displaced workers and to increase the federal minimum wage.

39

Prior to joining CBO, Todd worked as a Health Insurance Specialist at the Health Care Financing Administration and as a revenue analyst for a private-sector health care contractor. He holds a Masters degree in Public Administration from the University of Washington.

Eric Feuer, Ph.D., Eric J. "Rocky" Feuer, PhD, has been the Chief of the Statistical Research and Applications Branch since 1999. Prior to that he was head of the Surveillance Modeling and Methods Section in the Division of Cancer Prevention and Control. He has been at the NCI since 1987. Before coming to NCI, he was the Chief Statistician for the Cancer Center at Mount Sinai School of Medicine in New York.

Dr. Feuer received his Ph.D. in Biostatistics from the University of North Carolina at Chapel Hill in 1983. His research has focused on evaluating and developing new cancer progress measures; modeling the impact of cancer control interventions on the cancer burden; and developing statistical methods for the analysis, interpretation, and presentation of national cancer statistics. He is the author of over 75 peer-reviewed publications.

Dr. Feuer was elected as a fellow of the American Statistical Association in 2000. He received the NIH Director's Award in 1999 in recognition of his work with advancing statistical methods to interpret national cancer statistics. He serves on the Regional Committee of the Eastern North America Region (ENAR) of the International Biometric Society and was the program chair of its annual meeting in 1997. He was a statistical editor for the *Journal of the National Cancer Institute* from 1994 to 2000. He is a cofounding editor of the JNCI Cancer Surveillance Series, and serves as a reviewer for journals in statistics, epidemiology and cancer surveillance and control.

Karen M. Kuntz, ScD., is Associate Professor of Decision Science at the Harvard School of Public Health and is based at the Harvard Center for Risk Analysis. Dr. Kuntz's research focuses on the methodology and application of decision analysis and cost-effectiveness analysis in the evaluation of clinical and public health strategies. She is Principal Investigator of one of the NCI-funded Cancer Intervention and Surveillance Modeling Network (CISNET) grants to evaluate national trends in colorectal cancer incidence and mortality. She is also directing another NCI-funded project to develop a simulation model of colorectal and breast cancer that incorporates risk factors, screening, and treatment. Dr. Kuntz has also been involved in a number of other cancer-related disease models that evaluated health gains and/or cost-effectiveness associated with cancer prevention strategies as well as treatment and staging strategies for patients with cancer. In addition, Dr. Kuntz has extensive modeling experience in the clinical area of cardiovascular disease, conducting economic analyses on treatment and diagnostic approaches for patients with suspected carotid disease, patients after myocardial infarction, patients with stable chest pain, and primary prevention. In addition to specific applications, Dr. Kuntz has published on a number of methodological issues involving the evaluation of biases in decision modeling and the relationship between expert panel and decision-analytic outcomes. Recently, Dr. Kuntz was elected President Elect of the Society for Medical Decision Making and has assumed the role of Director for the AHRQ-funded Post-Doctoral Fellowship in Health Services Research at Harvard. Dr. Kuntz received her masters and doctorate, both in biostatistics, from the Harvard School of Public Health.

Uri Ladabaum, M.D., M.S., is currently Assistant Clinical Professor of Medicine in the Division of Gastroenterology at the University of California, San Francisco. Dr. Ladabaum completed undergraduate studies in Biochemistry at the University of California, Berkeley, in 1987 and graduated from the University of California, San Francisco School of Medicine in 1991. He served as a Resident in Internal Medicine from 1991 to 1994 and then as Chief Resident in Internal Medicine from 1994 to 1995 at Stanford University

Hospital. Dr. Ladabaum completed a three-year clinical and research Fellowship in Gastroenterology at the University of Michigan, Ann Arbor in 1998. During that time, he completed an M.S. program in Clinical Research Design and Statistical Analysis at the University of Michigan School of Public Health. Dr. Ladabaum served as Lecturer in the Division of Gastroenterology at the University of Michigan, Ann Arbor from 1998 to 1999. In 1999, Dr. Ladabaum joined the Division of Gastroenterology at the University of California, San Francisco as an Assistant Clinical Professor of Medicine and Director of the Gastrointestinal Motility Program.

Dr. Ladabaum's research focuses on two areas: health services research and cost-effectiveness analyses and the functional gastrointestinal disorders. In the first area, Dr. Ladabaum's research has centered on decision analytic models of colorectal cancer screening and chemoprevention in persons at average or above average risk for colorectal cancer and has included evaluations of management strategies for dyspepsia, screening for hepatocellular carcinoma, and evaluation and pharmacotherapy for irritable bowel syndrome. In the second area, Dr. Ladabaum has performed multiple studies of gastrointestinal physiology and visceral sensation. Colorectal cancer screening continues to be an active area of investigation for Dr. Ladabaum. His current efforts include decision analytic evaluations of emerging technologies for colorectal cancer screening and models of the potential population-wide clinical and economic impact of colorectal cancer screening in the United States.

Theodore R. Levin, M.D., is a Research Scientist II at the Kaiser Permanente Division of Research in Oakland, CA, and a Senior Physician, in the Gastroenterology Department of the Kaiser Permanente Walnut Creek Medical Center. He has been involved in colorectal cancer-related research for over 8 years. He serves on the American Cancer Society Colorectal Cancer Operations Committee and on the U.S. Multi-Society Task Force on Colorectal Cancer. He received his B.S. in Psychology from Duke University, an M.S. in Physiology from Georgetown University, and an M.D. from the Emory University School of Medicine. He did his internship and residency in Internal Medicine at University of Califorina at San Francisco (UCSF). He completed fellowships in Gastroenterology and Health Policy Studies also at UCSF, and continues to serve as an Assistant Clinical Professor of Medicine in the Gastroenterology Division at UCSF.

His research interests have focused on the clinical delivery of colorectal cancer screening. Specific interests are: the delivery of flexible sigmoidoscopy (ability of sigmoidoscopy to predict advanced proximal colonic neoplasia, complications of sigmoidoscopy, and the appropriate interval between screening sigmoidoscopy examinations), colonoscopy (complications of colonoscopy), and the use of novel screening tests such as immunochemical FOBTs and fecal DNA tests. He has served as PI of NCI RO1 grants evaluating the addition of immunochemical fecal occult blood tests to sigmoidoscopy screening and on the relationship between insulin resistance and adenomas of the colon and rectum. He is also currently the Kaiser Permanente PI of the NCI/Mayo Clinic multi-center study of fecal DNA screening.

Brian P. Mulhall, M.D., M.P.H., Dr. Mulhall is a gastroenterologist at Walter Reed Army Medical Center in Washington, DC. He obtained a B.S. in Biology and a B.A. in Philosophy/Ethics from the University of San Diego. He attended medical school at St. Louis University School of Medicine and graduated with a Distinction in Research. He completed his Residency in Internal Medicine at Madigan Army Medical Center, where he stayed as Chief of Medical Residents. He subsequently completed his Gastroenterology Fellowship at Walter Reed Army Medical Center while completing his M.P.H. at the Uniformed Services University of Health Sciences. He is now an Assistant Professor of

Medicine at the Uniformed Services University of the Health Sciences. He is active in several professional societies and is currently a member of the Educational Affairs Committee for the American College of Gastroenterology. He has multiple publications and abstracts in a range of topics, including nocturnal gastroesophageal reflux, infectious esophagitis, colorectal cancer screening, and fatty liver disease.

Currently, Dr. Mulhall is involved in research looking at rates of adherence to surveillance guidelines, along with the institutional and social-cognitive barriers that negatively impact adherence. He is also examining the utilization of resources surrounding colorectal cancer screening, with an emphasis on development of practices that will maximize outcomes while preserving resources. He had the opportunity to work with Dr. Joseph Lipscomb at the National Cancer Institute this past year on a project assessing the value on linear programming as a tool for modeling constraints in colorectal cancer screening—an eye-opening experience. Frankly, though, Dr. Mulhall spends most of his time with an endoscope in-hand striving to prevent colorectal cancer.

Reid Ness, M.D., M.P.H., is an Assistant Professor of Medicine at the Vanderbilt University Medical Center. He was recruited to Vanderbilt in 2000 after a health serves research fellowship at Indiana University where he also received a Masters in Public Health in Health Education. Dr. Ness's primary research interest focuses on quality improvement in the provision of colorectal cancer preventative services. He is actively engaged as a co-investigator in several projects. These include the (1) MMC/VICC Partnership (U54/NCI: Adunyah/Moses PIs) pilot project examining predictors of non-adherence with recommended colonoscopy, (2) Vanderbilt GI Cancer SPORE project 5 which seeks to examine predictors of colorectal adenoma recurrence in a prospective cohort (SPORE/NCI: Coffey PI), (3) NCI funded study to examine predictors of colorectal adenoma recurrence in a retrospective cohort (RO1/NCI: Zheng PI), (4) NCI funded project to create a health policy model of CRC for use in cost-effectiveness analysis (RO1/NCI/AHRQ: Dittus PI), and (5) Southern Community Cohort Study (RO1/NCI: Blot PI), a large prospective cohort study examining factors associated with cancer outcomes disparities between Caucasian and African Americans. He was also the principal investigator on a pilot project funded through the VICC (VICC: Ness PI) entitled the "Nashville Colorectal Health Study" the goal of which was to gather pilot data for predictors of colorectal adenoma identification at colonoscopy, cognitive/emotional outcomes of colonoscopy, and micro-costing for colonoscopy.

Michael P. Pignone, M.D., M.P.H., is an Assistant Professor of Medicine in the Division of General Internal Medicine at University of North Carolina-Chapel Hill (UNC). He received his medical degree and residency training in primary care internal medicine from the University of California at San Francisco. He then received fellowship training in clinical epidemiology and health services research through the Robert Wood Johnson Clinical Scholars program at UNC.

Dr. Pignone's research is focused on chronic disease prevention and physician- patient communication about risk in primary care settings. His main areas of interest include heart disease prevention, colorectal cancer screening, and disease management for common chronic illnesses such as diabetes, depression, heart failure, and chronic pain. He has conducted research examining the role of literacy in physician-patient communication and its effect on health outcomes, including racial/ethnic disparities

Deborah Schrag, M.D., M.P.H., is a health services researcher in the Department of Epidemiology and Biostatistics at Memorial Sloan-Kettering Cancer Center. She is also a medical oncologist in the Department of Medicine with a clinical practice dedicated to the treatment of patients with colorectal cancer. She is an Assistant Profesor of Medicine

and Public Health at the Weill Cornell School of Medicine. She received her M.D. from Columbia University College of Physicians and Surgeons and then completed a residency in Internal Medicine at the Brigham and Women's Hospital in Boston. She trained as a medical oncologist at the Dana-Farber Cancer Institute with a focus on treatment of gastrointestinal malignancy, specifically colorectal cancer. Following her clinical training, she earned an M.P.H. degree at the Harvard School of Public Health with emphasis on cancer-related decision-analytic modeling. Her current research interests focus on analyzing dissemination of new cancer treatment strategies and understanding the mechanisms that lead to disparities and variation in the quality of care. In particular, her research focuses on the utilization of population-based data resources such as SEER-Medicare to evaluate patterns and costs of care as well as the quality of care. Her research has demonstrated that cancer therapies that are well accepted as standard of care are not consistently delivered to definable subsets of vulnerable patients, particularly the elderly and members of racial and ethnic minority groups.

Laura C. Seeff, M.D., Dr. Seeff is a medical officer in the Division of Cancer Prevention and Control (DCPC) at the Centers for Disease Control and Prevention (CDC) in Atlanta, Georgia. She is the principal investigator for a CDC study to estimate the capacity of physicians and non-physician endoscopists across the country to provide endoscopic colorectal cancer screening to all eligible persons. Published results from this study will be available in 2004. She is also overseeing nine state-level colorectal cancer capacity assessments in Iowa, Michigan, Texas, Maryland, Massachusetts, Minnesota, New Mexico, New York and Washington state. Additional states will be added to this study in FY 05. She has authored several studies documenting low rates of use for colorectal cancer screening tests in the United States, and is currently overseeing a study evaluating the complication rate in colonoscopies performed in asymptomatic persons. She provides medical and scientific consultation for the multi-year federal colorectal cancer action campaign, *"Screen for Life"*, which is entering its fifth year. She helped develop an educational slide set for health care providers entitled "A Call to Action: Prevention and Early Detection of Colorectal cancer", which is available on the CDC website (www.cdc.gov/cancer/colorctl/calltoaction). Dr. Seeff also investigates cervical cancer screening issues among foreign-born women living in the United States. Prior to joining CDC in July of 1998, Dr. Seeff was an assistant professor in the Department of Medicine at the Emory University School of Medicine, where she was responsible for establishing a colorectal cancer screening clinic for patients at Grady Memorial Hospital, performing screening sigmoidoscopies and teaching sigmoidoscopy to doctors-in-training.

Marjolein van Ballegooijen, M.D., Ph.D., is an Assistant Professor of Medical Technology Assessment in the Screen Section of the Department of Public Health at the Erasmus MC (University Medical Center Rotterdam).

She received her medical degree from the University of Amsterdam, after which she joined the Department of Public Health in Rotterdam in 1987. In the first years, her main area of interest was cervical cancer screening. She worked on several national and international projects in this field, being responsible for the clinical and other medical input. This work resulted in her Ph.D., which she received Erasmus MC. The analysis performed for the national evaluation of program cervical cancer screening, had largely contributed to the updating of the official recommendations in 1994, e.g., the change in age-ranges and intervals for the program screening. Since 1996, she is involved in several national and international research projects on colon cancer screening.

Dr. van Ballegooijen's research is focused on decision-making concerning cancer screening, on basis of optimization of screening strategies from a cost-effectiveness point of view. In most of the projects she worked on, the MISCAN simulation program was used for the analysis of screening and epidemiologic data and the prediction of effects and costs.

Sally W. Vernon, M.A., Ph.D., Dr. Vernon is Professor of Behavioral Sciences and Epidemiology, Director of the Division of Health Promotion and Behavioral Sciences, and Senior Investigator in the Center for Health Promotion and Prevention Research, University of Texas-Houston, School of Public Health.

Her training is in epidemiology and behavioral sciences. Her interdisciplinary training has led to a research focus on translating epidemiologic evidence and principles to the design and evaluation of interventions to encourage the adoption of health behaviors. Her research interests encompass the epidemiology of mental disorders, with a particular focus on racial/ethnic differences in conceptualization and measurement; the analysis of psychosocial factors as predictors of health behaviors and health status indicators; and the design and evaluation of interventions to encourage adoption of cancer screening behaviors. Over the past 15 years, Dr. Vernon has published extensively in the area of cancer prevention and control with an emphasis on the primary and secondary prevention of breast, cervical, and colorectal cancers. More recently, she has worked in the area of informed consent and informed decision making about testing with the prostate specific antigen. In her recent work she has applied findings from epidemiologic studies to develop interventions to increase regular cancer screening behaviors, for cancers where the epidemiologic evidence supports the use of screening tests (e.g., cervical cancer) or to educate physicians and patients about the risks and benefits of tests with uncertain efficacy (e.g., prostate specific antigen).

Sandeep Vijan, M.D., M.S., is a Physician-Scientist at the Ann Arbor VA HSR&D Center of Excellence and an Assistant Professor of Internal Medicine at the University of Michigan Medical School. His research interests include evaluating the effectiveness and cost-effectiveness of preventive interventions in primary care, particularly those related to screening for common diseases and disease complications. He is also involved in evaluating methods of tailoring interventions based on individualized assessments of patient risks and preferences. Although he has a broad range of disease interests, his research has been primarily focused on diabetes and colorectal cancer.

Judith L. Wagner, Ph.D., Judith Wagner, Ph.D. has been a Scholar in Residence at the Institute of Medicine since January 2003. She has more than 30 years experience in health policy analysis and health technology economics. Most recently, as a Senior Analyst at the Congressional Budget Office, she specialized in prescription drug issues, including the design of a Medicare prescription drug benefit, Medicaid drug payment, and reform of current laws governing the entry of generic drugs into the market place. Before joining CBO, she was a consultant at the Mayo Clinic in Rochester, MN, where she conducted cost and cost-effectiveness analyses of medical procedures and technologies for both research and operational planning at the Clinic. She also managed major assessments of the cost-effectiveness of preventive and diagnostic technologies at the U.S. Congress Office of Technology Assessment, including studies of the cost-effectiveness of colorectal cancer screening in average-risk adults. Dr. Wagner holds a Ph.D. from Cornell University, where she studied economics and operations research with emphasis on environmental applications. She also holds master's degrees in economics from the University of Michigan and in environmental systems engineering from Cornell University.

Appendixes D-Q

Introduction

Appendixes D-Q contain the slide presentations of 14 speakers who were invited to address the workshop on specific topics. The presentations are reprinted here as they were presented to workshop participants, with little editing except format changes to permit printing and the addition of text summarizing accompanying remarks or clarification not shown on the slide. The original PowerPoint presentations are available on the Institute of Medicine website at http://www.iom.edu/crcworkshop.

Appendix D

Overview of Harvard Model
Karen M. Kuntz, ScD.

SLIDE 1

SLIDE 1 NOTES: The "Harvard Model" described here is the one presented in Frazier and colleagues (Frazier et al., 2000). That model was used to produce the results of the Pre-workshop Exercise described elsewhere in the workshop summary.

47

SLIDE 2

Model Structure

- Cohort model
  - follows a 50 yo cohort until death
- Markov model
  - 60 states of health
- From age 50 until death (or 85 for workshop analyses)
- SMLTREE software

2

SLIDE 2 NOTES: The model is a Markov design, programmed in SMLTREE™, a decision analysis software package. It follows a cohort of 50-year-old persons until death. (For the pre-workshop exercise, individuals were followed through age 85.) Sixty states of health track different underlying disease states, such as whether a person has a low-risk or high-risk polyp, and the location of that polyp, whether there is an undiagnosed or diagnosed cancer, and the stage of that diagnosis.

SLIDE 3

Modeling the Natural History of CRC
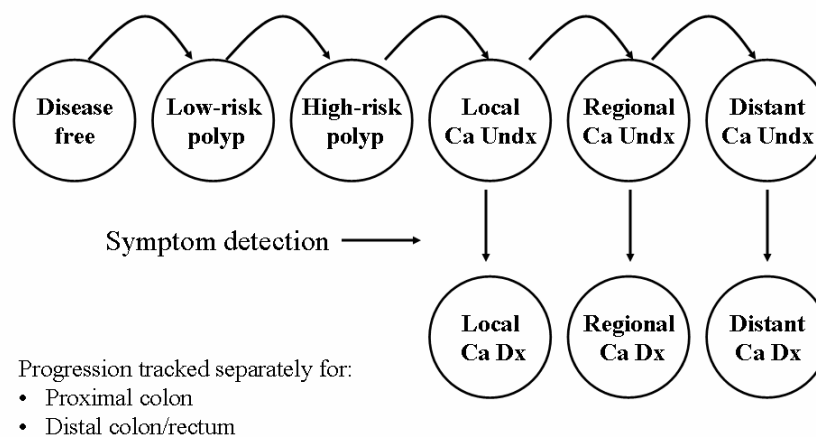
SLIDE 3 NOTES: This slide lays out a very simple diagram of the various states of health in which an individual can reside. Some states of health are free of any polyps or cancer in the colorectal tract. Each year there is a chance that a polyp will develop. Polyps then can progress to undiagnosed cancer. Cancers are staged using the NCI SEER nomenclature: localized, regional and distant. Every year, there is a chance that a cancer will be diagnosed as a result of symptoms. The probability of such detection varies with the stage of the patient.

The progression described above is tracked separately for the distal and proximal colon. Thus, the model can determine whether a particular polyp or cancer is in the part of the colorectal tract that is reachable by sigmoidoscopy. We assumed that all cancers arise from polyps, but not all polyps are destined to become cancers.

Definitions: Low-risk polyp=an adenomatous polyp smaller than 10 cm and with tubular histology. High-risk polyp=an adenoma greater than or equal to 1 cm or containing villous histology. Ca=cancer. Undx=undiagnosed. Dx=diagnosed.

SLIDE 4

Available Data

- Polyp prevalence from autopsy studies
  - Transition probabilities derived to match prevalence by age, sex, and location
- Screening studies
  - Proportion of polyps high risk
  - Transitions from low-risk to high-risk polyp
  - Transitions from high-risk polyp to localized cancer
- Estimates of national cancer incidence (SEER)
  - Transition probabilities through cancer stages and symptom detection by stage
- Survival curves by cancer stage (SEER)

4

SLIDE NOTES 4:   We estimated age- and sex-specific prevalence of adenomatous polyps using a weighted logistic regression analysis of results from 6 autopsy studies. We then derived transition probabilities to low-risk adenoma such that the model reproduced the age- and sex-specific prevalence.

Polyps were distributed across proximal colon, distal colon/rectum, or both sites based on data from autopsy studies and screening studies. We used screening studies to estimate the proportion of adenomas that are high risk.

The transition probability from low-risk to high-risk adenoma was estimated from studies of small polyps left in situ and reexamined annually. The transition probability from high-risk adenoma to invasive localized cancer was estimated from a study of patients who refused resection of high-risk polyp.

We varied the estimates of cancer progression and symptom detection across clinically plausible ranges so that the stage distribution and cancer incidence was similar to those reported by the National Cancer Institute's SEER system.

SLIDE 5

Autopsy Studies
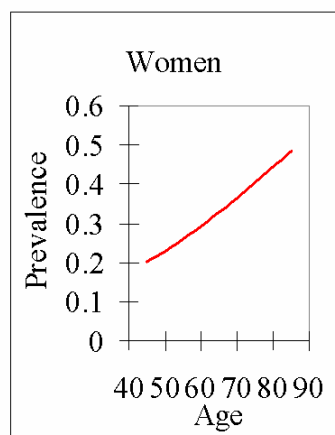
Women

Provide estimates of polyp prevalence by age and sex, which inform the probability of transitioning from disease free to polyp.

5

SLIDE 5 NOTES: This chart shows the fitted regression line for women. The model transition probabilities were constructed so that, for example, in the cohort of women who had survived to 60 years of age, roughly 27 or 28 percent would have at least one polyp.

SLIDE 6

Calibration of the Model

- Percent diagnosed with cancer in 10, 20, and 30 years, given cancer-free at age 50.

Years	SEER	Model
10	0.76%	0.66%
20	2.45%	2.76%
30	4.79%	5.17%
Lifetime	6.64%	6.69%

6

SLIDE 6 NOTES: We adjusted the transition probabilities from normal status to polyps, from polyps to undiagnosed cancer, and from undiagnosed to diagnosed caner until we achieved reasonable agreement with available data on cancer incidence. To do this, we used DevCan, a feature of the NCI SEER database which estimates the probability of developing cancer in one's lifetime. We used DevCan data from the late 1970's to predict cancer rates for a population that would not routinely receive screening. (Probabilities for recent years have most likely been affected by the rapid increase in the utilization of colorectal screening tests)

Here is an example of how our resulting model performed in predicting cancers in a cohort of men who were cancer free at age 50.

How close is close enough? At the time we published our paper, we used informal comparisons (i.e., the model predictions and the SEER predictions were "reasonably" close by visual inspection). Today, we are in the process of developing more sophisticated techniques, specifically statistical likelihood-based techniques to calibrate our models. Those improvements are not reflected in the modeling exercise we undertook for this workshop.

SLIDE 7

# Calibration of the Model (cont.)

- Stage distribution of CRC (lifetime for 50-year-old men).

Stage	SEER	Model
Localized	39%	40%
Regional	39%	38%
Distant	22%	22%

7

SLIDE 7 NOTES: We also calibrated the stage-distributions predicted in the model (in the absence of screening) with stage distributions reported in the NCI SEER database. We used data from the few studies that followed the progression of polyps not removed on first detection, estimates of the dwelling time of cancer by stage, estimates of symptom detection rates, to guide our initial model assumptions, and we ultimately adjusted those assumptions to reach reasonable agreement with the SEER data.

SLIDE 8
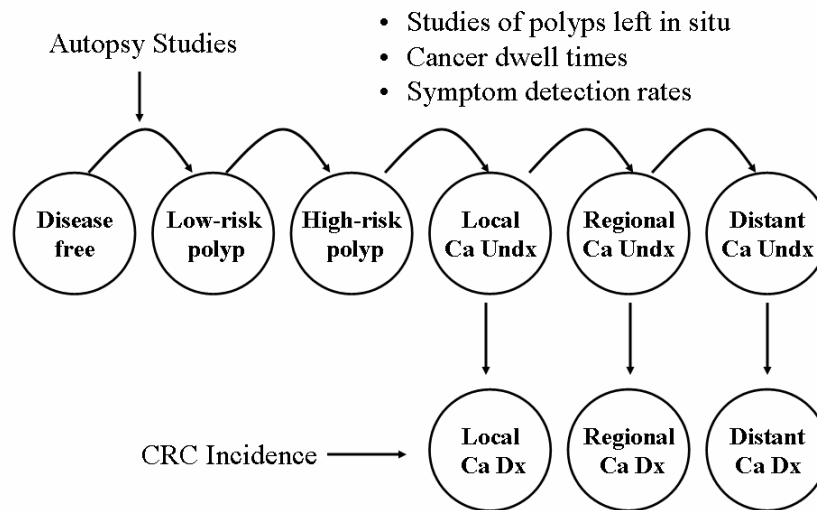
SLIDE 8 NOTES: This chart summarizes the sources of data used to predict the natural history of colorectal cancer in the absence of screening.

SLIDE 9

Assumptions About Compliance

- Compliance is a probability that gets applied each screening year and represents the percentage of persons who show up for their "scheduled" visit.

- Can be different for each initial test and for follow-up/surveillance colonoscopy.

- Is independent of past compliance history

9

SLIDE 9 NOTES: For example, if a screening strategy involves an annual FOBT test, and if we are assuming a 60 percent compliance rate, then each individual has a 60 percent chance of actually getting an FOBT in each year. The compliance in any particular year is independent of past compliance history.

We also assume that the compliance rate might differ for each screening technology and for follow-up colonoscopies. For example, the compliance rate could be 60 percent for FOBT, but 80 percent for colonoscopy scheduled following a positive screen.

SLIDE 10

Assumptions About Costs

- CRC costs
  - From Group Health Cooperative data
  - Depends on stage and varies with patient's remaining life
  - Initial, continuing, final
- Non-medical costs not included
- Extra medical care costs associated with living longer not included.

10

SLIDE 10 NOTES: The source of screening procedure and CRC treatment costs was the Group Health Cooperative of Puget Sound (Taplin et al., 1995a; Taplin et al., 1995b). That study reported on the net costs of the initial treatment for CRC, a monthly cost of continuing care, and a final cost associated with the last year in which the patient lives. Thus, the longer a patient lives, the more CRC treatment costs he or she will incur. Procedure costs were obtained from the same source through personal communication.

We did not include non-medical costs or the extra medical care costs associated with living longer.

SLIDE 11

CRC Screening
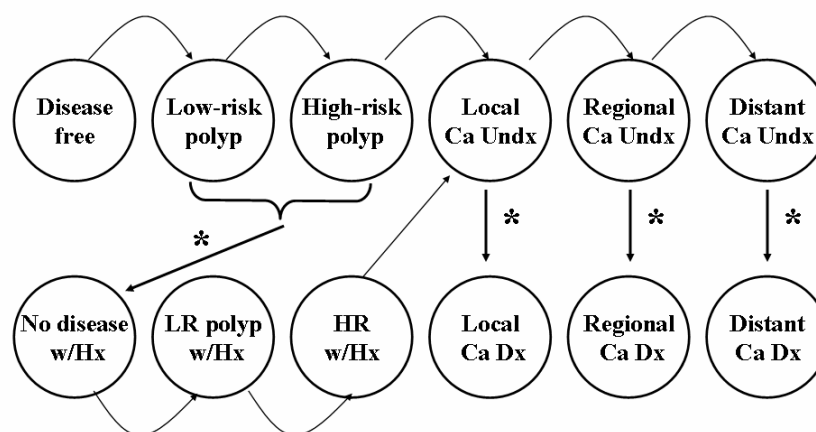
SLIDE 11 NOTES: To model a screening strategy, we superimposed a screening mechanism on the natural history model (the effects of screening are indicated by the asterisks). We modeled a screening, follow-up and surveillance strategy as follows: If a person has an underlying polyp and is screened, there is a chance (based on the sensitivity of the test) that the polyp will be found and removed. The person is now "disease free" but the model keeps track of whether that person had a low-risk or high-risk polyp (as indicated by "w/Hx"). Those with high-risk polyps can be entered into a surveillance protocol consisting of periodic colonoscopy. For the pre-workshop exercise, however, all individuals with polyps were assumed to undergo periodic surveillance. If a person has undiagnosed cancer and is screened, there is a chance that the cancer will become detected.

We assumed that recurrence rates of low-risk polyps after polypectomy were higher for individuals with a history of a high-risk polyp diagnosis compared with a prior diagnosis of low-risk polyp and that the risk was higher in the first year following polyp removal compared with subsequent years.

SLIDE 12

Model Validation:
Minnesota Trial

- Run a cohort of patients similar to trial cohort (age distribution, compliance) for 15 years
- Trial showed 33% reduction in CRC mortality and 11% reduction in incidence
- Model showed 31% reduction in CRC mortality and 7% reduction in incidence

12

SLIDE 12 NOTE: We validated the model against the results of the Minnesota trial of bi-annual FOBT (Mandel et al., 1999; Mandel et al., 2000). We entered a cohort of patients similar to those who participated in the trial and used the compliance rates observed in that trial. We ran our model for 15 years, the duration of the Minnesota trial. The screened individuals in the trial showed a 33 percent reduction in cancer mortality and an 11 percent reduction in cancer incidence. In comparison, our model showed a 31 percent reduction in cancer mortality and a 7 percent reduction in incidence. So, our model was reasonably consistent with the results of that trial.

SLIDE 13

Model Strengths and Weaknesses

- Cohort model is limits the flexibility of the model assumptions
  - E.g., cannot make test performance or compliance a function of patient history
- Current version of the model (work in progress)
  - Monte Carlo model
  - Incorporates several risk factors and their impact on the underlying progression of disease
  - More formal approach to calibrations

13

SLIDE 13 NOTES: The model presented here has certain limitations that restrict our flexibility in altering model assumptions. It is difficult with a Markov structure to maintain detailed information on each individual's history over time, because each additional piece of information requires another set of health states, and the number of health states quickly multiplies. For example, it would be prohibitively complex to make test performance or compliance a function of patient history.

We are currently testing a new version, which is a Monte Carlo design. In such models, it is easy to keep each patient's history as he or she ages. The new model also incorporates eight new risk factors which alters the underlying progression of disease.

Finally, we are taking a more formal approach to calibrating our model assumptions. We are using statistical maximum likelihood methods to optimize our natural history assumptions.

The new model is a precursor to one of several collaborating as part of the NCI's CISNET program, which is discussed in Dr. Rocky Feuer's presentation later in this workshop.

REFERENCES

Frazier AL, Colditz GA, Fuchs CS, Kuntz KM. 2000. Cost-effectiveness of screening for colorectal cancer in the general population. *JAMA.* 284(15): 1954–1961.

Mandel JS, Church TR, Bond JH, Ederer F, Geisser MS, Mongin SJ, Snover DC, Schuman LM. 2000. The effect of fecal occult-blood screening on the incidence of colorectal cancer. *N Engl J Med.* 343(22): 1603–1607.

Mandel JS, Church TR, Ederer F, Bond JH. 1999. Colorectal cancer mortality: Effectiveness of biennial screening for fecal occult blood. *J Natl Cancer Inst.* 91(5): 434–437.

Taplin SH, Barlow W, Urban N. 1995a. Erratum: Stage, age, comorbidity, and direct costs of colon, prostate, and breast cancer care. *J Natl Cancer Inst.* 87(8): 610.

Taplin SH, Barlow W, Urban N, Mandelson MT, Timlin DJ, Ichikawa L, Nefcy P. 1995b. Stage, age, comorbidity, and direct costs of colon, prostate, and breast cancer care. *J Natl Cancer Inst.* 87(6): 417–426.

Appendix E

Description of the Laudabaum
Colorectal Cancer Screening Model
Uri Ladabaum, M.D., M.S.

SLIDE 1

Appendix E
Description of the Laudabaum
Colorectal Cancer Screening Model

Uri Ladabaum, M.D., M.S.
Assistant Clinical Professor of Medicine
University of California, San Francisco

January 26-27, 2004

1

SLIDE 1 NOTES: This summary describes key elements of the current version of a model used in a cost-effectiveness analysis published by my colleagues and me (Song et al., 2004)

61

SLIDE 2

Model Structure

- Cohort Model
  - ➢ Average life-years and cost per person
- Can be aggregate annual model
- Markov model
- From age 50 to 85 years or death for current workshop
  - ➢ Can be until age 100 years or death
- Can be sex-specific
- Software: DataPro

2

SLIDE 2 NOTES: Our model is a cohort model. The main outputs are average years of life lived and accrued costs per person.

It can be an converted to an aggregate annual model by combining the estimates for every age group in each year and projecting to the national population. Year-2000 U.S. Census data can be combined with age-specific model outputs to make predictions for the U.S. population (aggregate annual model), as opposed to a hypothetical cohort of a given size starting at age 50 years.

The model is a Markov model. For the purposes of the pre-workshop modeling exercise, it followed people from age 50 to 85 years of age, or until death if that came before age 85. The model can incorporate stopping ages up to age 100. It is also possible to treat each sex separately, though most of our work has been with average values for the entire population.

The current version runs on DataPro™, a commercial software package.

SLIDE 3

Schematic of Natural History Model
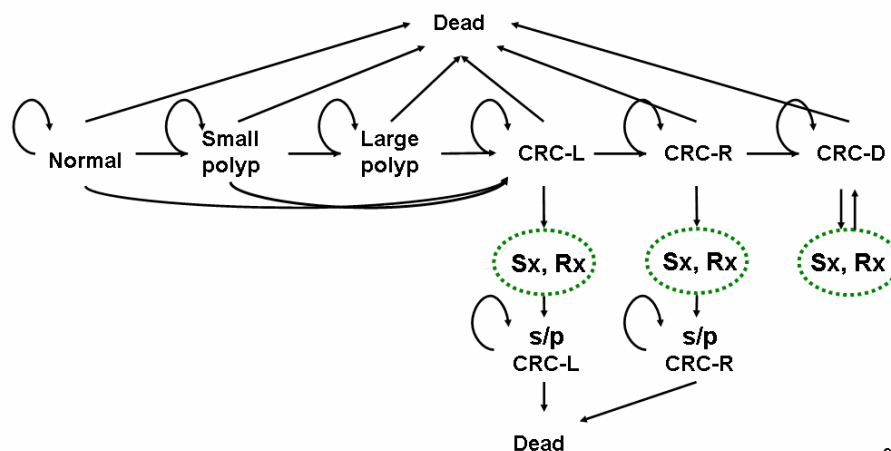
SLIDE 3 NOTES: This is a schematic of the natural history model. At the left, we start with people in the normal state. They can progress to a small polyp (less than 10 mm), which is our nomenclature for a low-risk polyp. Over time small polyps can progress to large polyps (greater than or equal to 10 mm) and then to localized, regional, or disseminated colorectal cancer.

    Some patients can progress directly from the normal state to localized colorectal cancer. In our model, approximately 85 percent of colorectal cancers arise from the adenoma-to-carcinoma sequence. The remaining 15 percent of cancers arise de-novo.

Our model posits that in the absence of screening the only way a cancer can be detected is through the emergence of symptoms. If patients present with symptoms, they are assumed to get the appropriate diagnostic workup and the appropriate therapy. If they survive, they enter a new state of "history of cancer."

    From each state of cancer patients can die. The probability of death was estimated from epidemiological data on stage-specific survival for the average population.

    Definitions: CRC-L=colorectal cancer, localized. CRC-R=colorectal cancer, regional. CRC-D=colorectal cancer, distant. Sx=symptoms. Rx=treatment s/p=status-post.

SLIDE 4

Approach to Modeling Natural History

- Polyp prevalence by age from autopsy studies
  - ➤ Small (<10 mm) *vs.* large (≥10 mm)
- Age and stage-specific cancer incidence from SEER
- 85% of cancers assumed to arise from polyps
- Localized to regional to disseminated cancer
  - ➤ Dwell time 2 years in each localized and regional
- Transition probabilities derived
  - ➤ Normal to small polyp: to yield autopsy rates
  - ➤ Small to large polyp: to yield autopsy rates
  - ➤ Large polyp to cancer, "mucosa" to cancer, symptomatic with cancer: to yield SEER by stage

4

SLIDE NOTES 4: Our model starts with the adenoma-carcinoma sequence. We model adenomatous polyps, which are generally agreed to be the precursor lesions for most colorectal cancers.

We started with polyp prevalence by age from onset (50 years old). We assumed total polyp prevalence at age 50 (15%, average of men and women; 5% of polyps are "large"). We derived transition probabilities from small to large polyps from data on prevalence at various ages.

We used age- and stage- specific cancer incidence rates reported in SEER (1990-1994). We assumed that 85% of cancers arise from polyps. We derived age-specific transition probabilities for normal to localized cancer (same for small polyp to localized cancer) and a fixed annual transition probability from large polyp to localized cancer. (It turns out that a fixed transition rate fits the data well, which is biologically plausible.)

We assumed a dwelling time of 2 years each in localized and regional cancer, rates of symptomatic presentation derived to match SEER stage distribution (22%/yr for localized and 40%/yr for regional; 100% for distant).

Stage-specific yearly mortality rates were derived for first five years after diagnosis (1.74%/yr for localized and 8.6%/yr for regional). For patients who survive more than 5 years, we applied age-specific mortality from all causes. We assumed distant cancer has average survival of 1.9 years.

SLIDE 5

Model *vs.* Polyp Prevalence (autopsy)
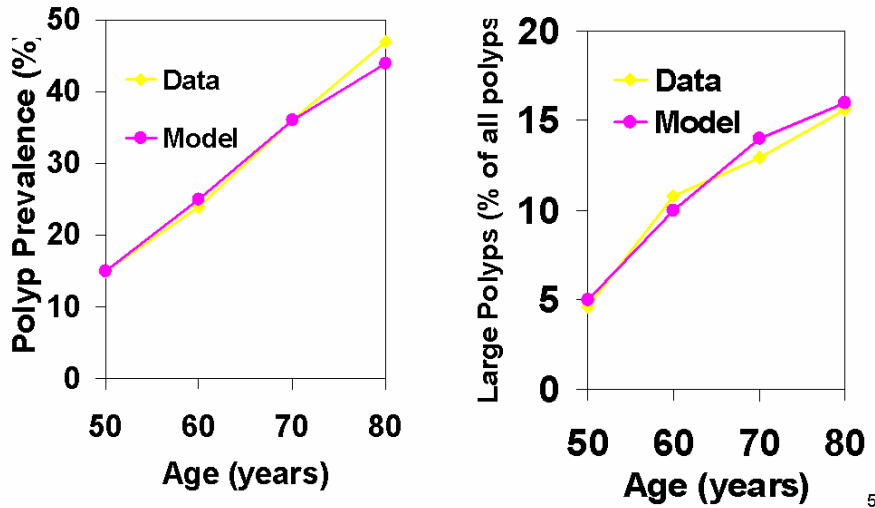
SLIDE 5 NOTES: To estimate the transition probabilities we used various data sources. For transitions from normal mucosa to small polyp, and from small polyps to large polyps we used age-specific data from autopsy studies (Rickert et al., 1979; Arminski and McLean, 1964; Williams et al., 1982; Vatn and Stalsberg, 1982; Clark et al., 1985).

The graphs above show how the predictions from our model so calibrated compare to the average of the published data on polyp prevalence for men and women.

SLIDE 6
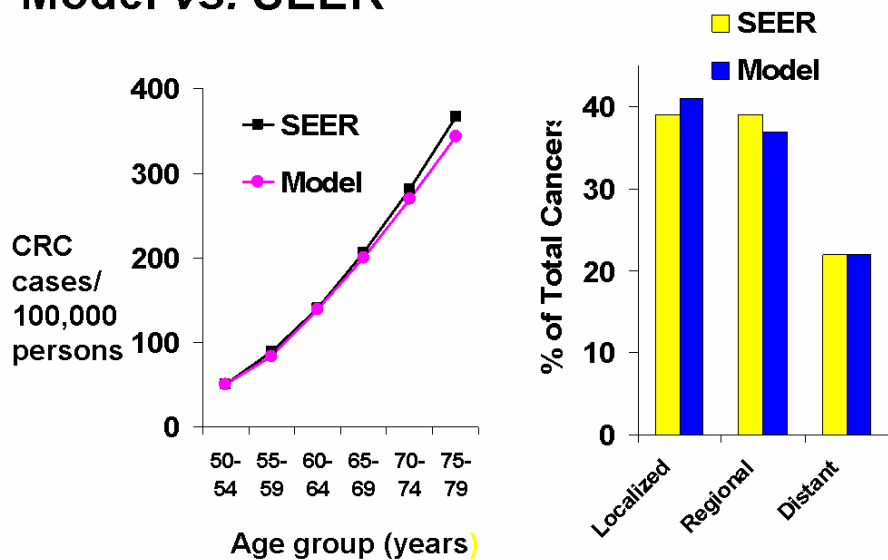
SLIDE 6 NOTES: The graphs in this slide show how the model, as calibrated by SEER data (Ries et al., 1997), compares to the age- and stage-specific incidence rates.

SLIDE 7

Assumptions About Compliance

- Persons either comply or do not comply with screening, follow-up, treatment and surveillance
  - ➢ *e.g.* If "25% compliance," 25% comply fully and 75% experience the "natural history"
- Base case: Assume perfect compliance with all strategies
- For any compliance rate <100%:
  - ➢ ICE ratios and strategy ranks *remain the same* as base case if compliance equal in all strategies
  - ➢ Results affected by differential compliance between strategies

7

SLIDE 7 NOTES: Incremental effectiveness of screening compared with no screening is affected by compliance, but incremental cost-effectiveness is not.

Note that when we assume less than 100 percent compliance, the incremental cost-effectiveness ratios and the rankings across screening strategies remain the same, provided that compliance is equal across all strategies. However, the incremental cost-effectiveness ratios and rankings across strategies would be affected if compliance rates vary across strategies.

Definitions: ICE = incremental cost-effectiveness.

SLIDE 8

Assumptions About Cost

- Costs of cancer care modeled as stage-specific "lump sum" on the year of diagnosis (discounted, like other costs)
    - ➢ Not dependent on age or life-expectancy
- Non-medical costs not included
- Costs of living longer if cancer prevented or cured not included

8

SLIDE 8 NOTES: Until now, the health care costs associated with dying of other causes has been set at 0. The model could accommodate other assumptions, however. In addition, those costs could be age-dependent, if reasonable data were available.

SLIDE 9

Model Validation: Natural History

- Predictions of Natural History model for U.S. using 2000 census data are consistent with published data:

	Total cases/yr	Total cancer care cost/yr
Model	159,000	$5.2 bil (yr 2003 $)
Published	148,000	$5.2 bil (yr 2000 $)

9

SLIDE 9 NOTES: To validate the model, we examined the age-specific outcomes of the model and compared them with national data for the year 2000. We were gratified that the estimated number of cancer cases in our model is consistent with published data (Jemal et al., 2003; Sandler et al., 2002).

SLIDE 10

Assumptions About Cost

- Costs of cancer care modeled as stage-specific "lump sum" on the year of diagnosis (discounted, like other costs)

  ➢ Not dependent on age or life-expectancy

- Non-medical costs not included

- Costs of living longer if cancer prevented or cured not included

1

SLIDE 10 NOTES: In the Minnesota study, screening occurred in only some of the years over the course of the trial, and imperfect adherence with screening was attained (Mandel et al., 2000; Mandel et al., 1993). Overall, in that study, about 50 percent of all the potential yearly screenings were actually completed. With full adherence with yearly FOBT, our model predicts approximately double the reductions in cancer incidence and mortality that were observed in the Minnesota trial.

For sigmoidoscopy every 5 years, our model predicts a 56 percent reduction in colorectal cancer incidence. That result is consistent with the Kaiser data for left-side lesions (Selby et al., 1992). The model's estimate is a bit higher, however, because in the model we account for a reduction in CRC incidence due to removal of polyps in the proximal colon at the time of follow-up colonoscopy (subsequent to a positive sigmoidoscopy In the Kaiser study, the benefit of sigmoidoscopy was limited to the part of the colon examined during screening.

The model predicted a reduction of 71 percent in cancer incidence for colonoscopy screening every 10 years. That result is consistent with findings of the National Polyp Study (Winawer et al., 1993)

SLIDE 11

Model Strengths and Weaknesses

- Model different levels of polyp and cancer risk

- Predictions at the level of the entire population

- Adjunct interventions evaluated (e.g. chemoprevention)

- Limitations and potential biases at present:

  ➢ Sequential/repeated test results independent

  ➢ Fixed cancer dwell time and % of cancers from polyp

  ➢ Polyp surveillance not dependent on size of polyp

  ➢ "Most advanced lesion" modeled (polyp by size)

  ➢ Location/histology not modeled ("FS reaches 50%")

11

SLIDE 11 NOTES: It is possible for the model to evaluate the costs and effectiveness of screening in populations with different levels of risk for polyps and cancer. We are currently working on making predictions at the level of the entire population. We have evaluated adjunct interventions, namely chemoprevention withy aspirin in average-risk patients and Cox-2 inhibitors in average-risk and high-risk patients (Ladabaum et al., 2001; Ladabaum et al., 2003).

The model does have certain limitations and potential biases. In addition to those listed, we are currently not satisfied with our ability to model complex patterns of compliance with screening, follow-up and surveillance. We are working to improve the model in that regard.

The question of independence between sequential tests (e.g. FOBT and FS) or repeated tests (e.g. annual FOBT in a cancer that is "dwelling") is probably important to predictions from the model (e.g. independent annual FOBT at 40% sensitivity for cancer has only a 0.6x0.6x0.6x0.6 = 13% chance of not picking up a cancer at some point before it is distant if cancer dwells in localized x 2 years and regional x 2 years)

As the model is currently structured, transitions among most patient states are probabilistic. However, for individuals with cancer, we assumed a fixed dwelling time in each stage. And the percent of cancers that arise from polyps has been fixed at 85 percent.

We are not currently modeling the details of polyp location (distal vs. proximal) or histology (high-risk vs. low-risk). However, we do assume that sigmoidoscopy can

reach 50 percent of all lesions. So, that is one strategy which might be evaluated differently if our model were to identify polyps by location.

REFERENCES

Arminski TC, McLean MD. 1964. Incidence and distribution of adenomatous polyps of the colon and rectum based on 1,000 autopsy examinations. *Dis Colon Rectum.* 7: 249–261.

Clark JC, Collan Y, Eide TJ, Esteve J, Ewen S, Gibbs NM, et al. 1985. Prevalence of polyps in an autopsy series from areas with varying incidence of large-bowel cancer. *Int J Cancer.* 36(2): 179–186.

Jemal A, Murray T, Samuels A, Ghafoor A, Ward E, Thun MJ. 2003. Cancer statistics, 2003. *CA Cancer J Clin.* 53(1): 5–26.

Ladabaum U, Chopra CL, Huang G, Scheiman JM, Chernew ME, Fendrick AM. 2001. Aspirin as an adjunct to screening for prevention of sporadic colorectal cancer: A cost-effectiveness analysis. *Ann Intern Med.* 135(9): 769–781.

Ladabaum U, Scheiman JM, Fendrick AM. 2003. Potential effect of cyclooxygenase-2 - Specific inhibitors on the prevention of colorectal cancer: A cost-effectiveness analysis. *Am J Med.* 114(7): 546–554.

Mandel JS, Bond JH, Church TR, Snover DC, Bradley GM, Schuman LM, Ederer F. 1993. Reducing mortality from colorectal cancer by screening for fecal occult blood. *N Engl J Med.* 328(19): 1365–1371.

Mandel JS, Church TR, Bond JH, Ederer F, Geisser MS, Mongin SJ, Snover DC, Schuman LM. 2000. The effect of fecal occult-blood screening on the incidence of colorectal cancer. *N Engl J Med.* 343(22): 1603–1607.

Rickert RR, Auerbach O, Garfinkel L , Hammond EC, Frasca JM. 1979. Adenomatous lesions of the large bowel: an autopsy survey. *Cancer.* 43(5):1847–1857.

Ries LAG, Kosary CL, Hankey BF, Miller BA, Harras A, Edwards BK. 1997. *SEER Cancer Statistics Review, 1973-1994.* Bethesda, MD: National Cancer Institute.

Sandler RS, Everhart JE, Donowitz M, Adams E, Cronin K, Goodman C, Gemmen E, Shah S, Avdic A, Rubin R. 2002. The burden of selected digestive diseases in the United States. *Gastroenterol.* 122(5): 1500–1511.

Selby JV, Friedman GD, Quesenberry Jr CP, Weiss NS. 1992. A case-control study of screening sigmoidoscopy and mortality from colorectal cancer. *N Engl J Med.* 326(10): 653–657.

Song K, Fendrick AM, Ladabaum U. 2004. Fecal DNA Testing Compared with Conventional Colorectal Cancer Screening Methods: A Decision Analysis. *Gastroenterol.* 126(5): 1270–1279.

Vatn MH, Stalsberg H. 1982. The prevalence of polyps of the large intestine in Oslo: An autopsy study. *Cancer.* 49(4): 819–825.

Williams AR , Balasooriya BA, Day DW. 1982. Polyps and cancer of the large bowel: a necropsy study in Liverpool. *Gut* 23(10): 835–842

Winawer SJ, Zauber AG, May Nah Ho, O'Brien MJ, Gottlieb LS, Sternberg SS, Waye JD, Schapiro M, Bond JH, Panish JF, Ackroyd F, Shike M, Kurtz RC, Hornsby-Lewis L, Gerdes H, Stewart ET, Lightdale CJ, Edelman M, Fleisher M. 1993. Prevention of colorectal cancer by colonoscopic polypectomy. *N Engl J Med.* 329(27): 1977–1981.

Appendix F

MISCAN-colon: An Overview
Marjolein van Ballegooijen, Iris Vogelaar, Rob Boer, Franka Loeve, Ann Zauber, Gerrit van Oortmarssen, and Dik Habbema

SLIDE 1

**Appendix F
MISCAN-colon: an Overview**

IOM meeting, 26th January 2004

Marjolein van Ballegooijen, Iris Vogelaar, Rob Boer,
Franka Loeve, Ann Zauber, Gerrit van Oortmarssen,
Dik Habbema

Dept. Public Health
Erasmus MC, Rotterdam

1

SLIDE 1 NOTES: No notes.

73

SLIDE 2

Model Structure

- Micro-simulation model programmed in Delphi

- Time-event model
- Discrete event – continuous time simulation

- "Parallel universe" model

- Real population-based model (published CEA)

- Flexible study-period and age-ranges → birth cohort model
  (IOM)

2

SLIDE 2 NOTES: Elements of the Miscan Model's structure are shown in this frame. The micro-simulation model treats time as a continuous variable, with discrete events occurring along the time line. So, for example, the appearance of a polyp of a particular size is a discrete event. (The model identifies three sizes.) But the dwelling time of such a polyp before it progresses to a larger size can be any length. Probability functions determine when the transition from one state to another will occur.

By parallel universe, we mean that the model first generates a complete simulated, or hypothetical, population of individuals with their complete natural history in the absence of screening. The results of that baseline simulation, with all pertinent clinical details, are stored. Then, the model begins again with the same population as before, but this time a particular screening strategy is imposed. Assumptions about the effect of screening on the clinical course of each patient determine a new set of simulated results. The net effect of the screening strategy is determined per life history by the changes that occur compared with the baseline.

The model provides outputs on a real population in any specific calendar year. Most of our published work has taken that approach. However, the model is flexible in that a specific cohort of individuals (such as 50 year old men) can be followed throughout the rest of their lives. For the pre-workshop exercise, we did adapt the model to a cohort structure.

SLIDE 3

Modeling Natural History

- Multi-disease model:
  - age-specific incidence rate combined with individual risk factor for disease onset
  - adenomas can arise at different places in the colon at different times and develop independently of each other within one person
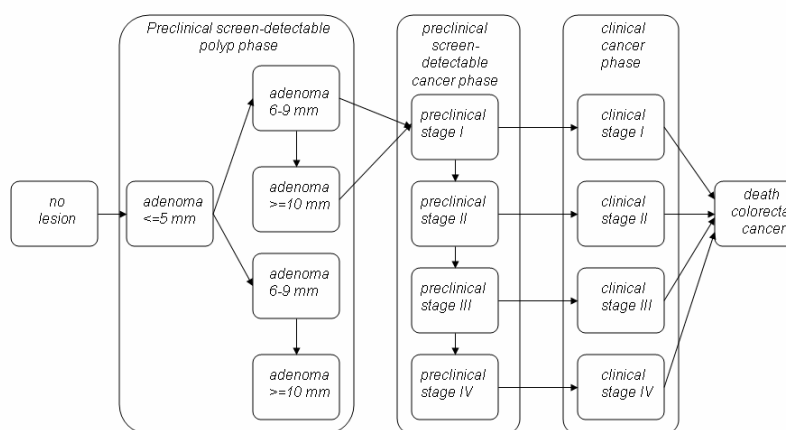
example:

onset of non-progressive adenoma in rectum

onset of progressive adenoma in transversum

3
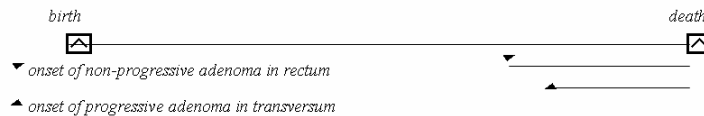
SLIDE 3 NOTES: The model is also a multi-disease model. By that I mean two things. First, although the model is intended to simulate a general population, it recognizes the inherent variation among individuals in the risk of developing adenomas and cancer. Thus, each individual has his own probability of acquiring one more adenomas.

Second, adenomas are simulated as distinct occurrences in the same person. One or more can occur simultaneously or over time. For example, a person's simulated natural history might be programmed to develop a non-progressive adenoma in the rectum at one time, and another progressive adenoma in the transverse colon at another time. Each of these events will have their own independent history until a first cancer is detected in that simulated individual.

SLIDE 4

Natural History: Adenoma Carcinoma Sequence

SLIDE NOTES 4: This chart shows the disease states and transitions that are included in the model. For example, a person with no adenomatous lesions may develop a small adenoma, which may or may not progress into cancer.
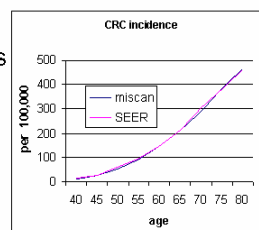
The length of time required for each person to travel through each phase in the adenoma/carcinoma sequence is determined by the parameters of a probability distribution for each disease stage and for each transition that is specified as part of the model.

The model as currently quantified does not permit a cancer to derive directly from a small adenoma. The adenoma must first progress to a medium or large size. Once the transition to cancer occurs, the lesion progresses through four stages. In any stage there is a chance, based on an assumed probability of detection, that the lesion will be diagnosed and become a clinical case. Once clinical, the lesion will or will not cause the patient's death from colorectal cancer with a specific probability specific to each stage.

SLIDE 5

SLIDE 5 NOTES: The parameters used to generate each hypothetical person's history were derived from a variety of sources, including data produced in clinical and epidemiological studies and, when no reliable data were available, expert opinion.

We started with data on the US population by age and calendar year. We also used SEER CRC incidence data from the late 1970's before screening was prevalent in the USA. We also used the extant autopsy studies to estimate adenoma prevalence, including the prevalence of multiple adenomas in the same individual.

SLIDE 6

Natural History: Quantification of the Parameters (2)

Direct estimates	Assumptions	Resulting from fit procedure
Demography	Duration distribution in preclinical states	Probability for an adenoma to be progressive
Distribution of lesions over large bowel	Transition probabilities from preclinical non-invasive states	Individual risk index
Survival after clinical diagnosis	Correlation between durations in subsequent states	Incidence rate of adenomas
Sensitivity, specificity and reach of screening tests	Dependency of test outcomes	Transition probabilities from preclinical invasive states
	Survival after screen-detected diagnosis	

6

SLIDE 6 NOTES: This chart provides an overview of which parameters were based on direct estimates and which resulted from a statistical fit procedure. Those based on assumptions are the ones that we must investigate when we validate the model, because they are the ones for which there are inadequate data from screening studies at present. They include the duration and distribution of the preclinical state and, closely connected, the transition probabilities from one preclinical state to another.

We assumed that it takes, on average, 20 years from the start of a small adenoma to a clinical cancer. We have attempted to validate this assumption with adenoma endoscopy studies, but more such studies are needed for us to be fully confident in that assumption. Another assumption regards the correlation among durations in different preclinical disease states. For now, we assume that if an individual progresses rapidly through the small adenoma state, he or she has a high probability of progressing rapidly through later states. We also assume a positive correlation among the durations of preclinical cancerous states.

At present, the results of screening tests are assumed to be independent of earlier or later tests, but we do not have data to support or reject this assumption. We also assumed that the same stage-specific survival holds for cancers detected through a screening procedure as for those detected through presentation of clinical symptoms. Thus, all the gain from early detection of cancer through screening results from a shift in the stages at which cancers are detected. Better data are needed to determine whether this assumption is justified.

SLIDE 7

Compliance

SCREENING COMPLIANCE
- Data considered:
  - Trials
  - Population surveys (NHIS)

- Compliance depends on prior compliance
  - Non-first screening ages: distinction in attenders (previous round) and non-attenders attendance

DIAGNOSTIC FOLLOW-UP AND POST ADENOMA-DETECTION
    SURVEILLANCE: fixed to 100%

7

SLIDE 7 NOTES: To arrive at realistic assumptions about the degree of compliance that can be expected from a particular screening strategy, we relied on data reported from clinical trials and population surveys. Population surveys are problematic as a source of compliance estimates, because the population has a choice of 3 or more different screening technologies for colorectal cancer screening, and the surveys do not ALWAYS provide detailed information on the specific test received. Applying compliance rates found in surveys to a screening strategy involving only one procedure (e.g., sigmoidoscopy every 5 years) could result in errors.

   Our model assumes that the probability of compliance with the screening strategy at any screening round depends on the person's compliance history in the previous screening round. Compliance in one round implies, in the model, a higher probability of compliance in the successive round. The probability of attending screening when attended previously is four times the probability when not attended the previous round. The total percentage of people who are attenders is constant over time.

   Compliance with all diagnostic follow-up and post-adenoma detection surveillance is assumed to be 100 percent.

SLIDE 8

Cost Assumptions

- Initial CRC treatment costs (first six months)

- Costs for continued (CRC related) care in the life years after CRC diagnosis

- Costs for terminal care (last six months) [not included in IOM runs]

- Non-medical costs are not included

8

SLIDE 8 NOTES: Our assumptions about the costs of screening, follow-up and surveillance procedures are shown elsewhere in the workshop summary. The costs of CRC treatment vary with the number of months since the detection of a CRC. Initial treatment costs cover the first six months; after that, continued CRC-related costs are charged over the remainder of the life, until the last six months, when costs of terminal care are included.

However, for the pre-workshop exercise, because individuals would be followed only until age 85, we did not include terminal care costs, even when a simulated person died before the age of 85.

We include only those costs associated with treating CRC. Unrelated medical care costs are not included, nor are any non-medical costs, such as transportation or work losses.

SLIDE 9

Model Validation

- Model validated on:
  - National Polyp Study
  - Kaiser Flexible Sigmoidoscopy study
  - Minnesota-trial (not finished)
- Working on further (*combined)* validation:
  - Funen-trial
  - Nottingham-trial
  - Minnesota-trial

9

SLIDE 9 NOTES: We have attempted to validate the model's predictions by comparing them with the results of several screening and surveillance trials. In comparison to the National Polyp Study (a trial of post-polypectomy surveillance) Miscan predicted more cancers in the years following polypectomy than was reported in the trial (Loeve et al., 2004). In comparison to the Kaiser Flexible sigmoidoscopy study (Levin and Palitz, 2002) we also found too many cancers during the 5 years following a negative sigmoidoscopy (van Ballegooijen et al., 2002). So, we predicted too many cancers during follow up in individuals both with and without adenomas, while our total risk for cancer was calibrated to the total population. To explain the sigmoidoscopy data, we need to adjust our model so that risk is shifted away from individuals with few or now adenomas. Consequently, individuals with adenomas (as in the National Polyp Study) on average will have an increased risk. But this is not compatible with the National Polyp Study, where we already simulated too many cancers. On the other hand, since neither of these two studies was randomized before entry, it cannot be ruled out that the study groups consist of relatively low-risk individuals through self-selection (in the sigmoidoscopy study) or through clinical selection (in the NPS). This shows why we badly need randomized studies for further validation.

We have been able to reproduce the results of the Minnesota trial. Our model predicted a 34.6 percent reduction in morality from annual screening and a 20 percent reduction from biennial screening. In the Minnesota trial the observed reductions were 33 percent and 21 percent, respectively. Although we reproduce the mortality reduction found in that trial, the model did not predict the stage distribution in follow-up rounds so

well. The model predicted a much more favorable stage distribution than was found in the Minnesota trial. That failure may suggest problems if we are to use the model for improved tests or other strategies. We are therefore currently attempting to validate the model against the three big FOBT trials together, combining the data that are available from each one.

SLIDE 10

Strengths and Weaknesses

- Strengths:
  - Multi-disease (multiple adenoma histories are simulated individually within one life history)
  - Both cohort and real population based approach possible
  - Flexible
- Weaknesses:
  - Fixed compliance to diagnostic follow-up and surveillance
  - Risk factors not included
  - Survival static over time
  - Flexibility causes complexity for users
- More validation …

10

SLIDE 10 NOTES: No notes.

SLIDE 11

Work in process

- Further calibration and validation
- Period- or treatment-dependent survival
- Adding risk factors
- Age-specific incidence per (risk)-stratum
- Expanding the screening module to reflect (historical) screening compliance to 3 different tests
- Define advanced adenoma states based on histology in addition to size (future)

11

SLIDE 11 NOTES: No notes.

REFERENCES

Levin TR, Palitz AM. 2002. Flexible sigmoidoscopy: An important screening option for average-risk individuals. *Gastrointest Endosc Clin N Am.* 12(1): 23–40.
Loeve F, Boer R, Zauber AG, Van Ballegooijen M, Van Oortmarssen GJ, Winawer SJ, Habbema JDF. 2004. National Polyp Study data: Evidence for regression of adenomas. *Int J Cancer.* 111(4): 633–639.
van Ballegooijen M, et al. 2002. *Validation Study of the Miscan-Colon Program Sith Data From the CoCaP Sigmoidoscopy Program. Final Report to NCI.* Rotterdam, NE: Erasmus Medical Center.

Appendix G

The Vanderbilt Colorectal Cancer Model
R.M. Ness, R.W. Klein, R.S. Dittus

SLIDE 1

Appendix G
The Vanderbilt Colorectal Cancer
Model

R.M. Ness[1], R.W. Klein[2], R.S. Dittus[1]
[1]*Vanderbilt University Medical Center*
[2] *Medical Decision Modeling, Inc.*

1

SLIDE 1 NOTES: The materials presented in this workshop are based on our current
working version of the model. It was adapted from a paper published in 2000 (Ness et al.,
2000).

84

SLIDE 2

Model Structure

- Discrete-event simulation model
- Model can be employed to perform cohort or annual aggregate simulation
- Capable of simulating patients of any age over any given time period
- Outputs: distributions of aggregate costs and life-years and/or QALYs
- Programmed in Insight 5.4

2

SLIDE 2 NOTES: Discrete-event simulation employs a computer modeling technique in which system evolution over time is represented by variables that change instantaneously at discrete points in time called "events" (e.g. adenoma incidence). The model moves forward through time by advancing from one event to the next. A discrete-event structure gives the modeler the ability to assign specific attributes to each simulated patient (e.g. the underlying risk of adenoma incidence).

The model is programmed in Insight, a FORTRAN-based programming simulation language.
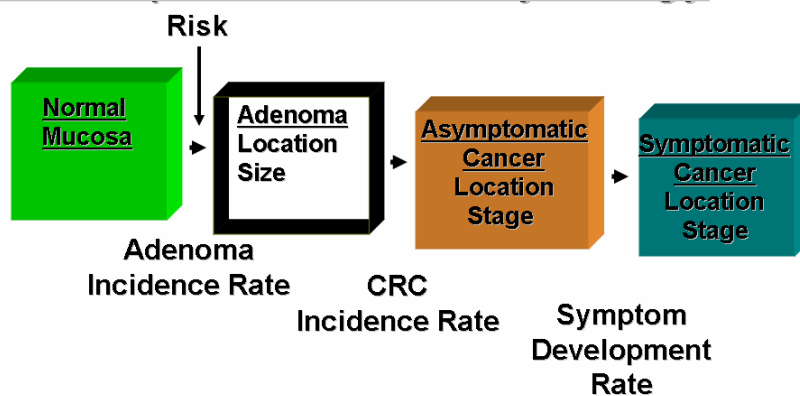
SLIDE 3

SLIDE 3 NOTES: All colorectal cancers (CRCs) originate only from pre-existing adenomas (this is a fair assumption for a U.S. population where the development of CRC from flat lesions has rarely been described).

The major steps in neoplasm development include normal tissue, adenoma, asymptomatic CRC and symptomatic CRC.

The variables that define each step in neoplasm development include location, size and stage.

The underlying genetic risk of sporadic (non-familial) CRC principally controls the adenoma incidence rate. In this model (as in the actual US population it simulates), each person has a differential risk of developing colorectal cancer in their lifetime. We used that risk distribution to predict the adenoma incidence.
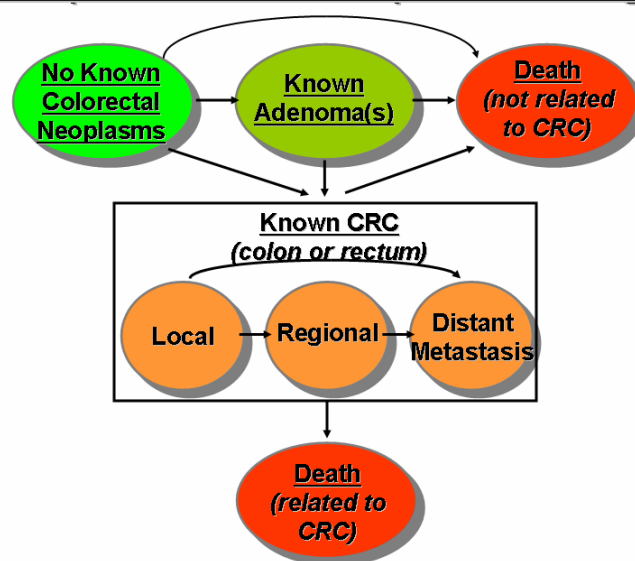
SLIDE 4

SLIDE NOTES 4: The health states that colorectal neoplasms can create include the no known adenoma, the known adenoma, the known CRC, the treated CRC and the death states.

The CRC dependent health states are defined by what was clinically known about the number and nature of any underlying colorectal neoplasms.

Anything that alters clinical knowledge such as diagnostic tests or surgical interventions can lead to transitions between the adenoma and CRC states.

SLIDE 5

Natural History Assumptions

- **All CRCs develop from pre-existing adenomas**
- **Less than 2.5% of adenomas will become CRC within the first 10 years**

- **There are adenomas which progress to CRC relatively quickly and those which progress only very slowly (dichotomous population)**

- **Dwell times fitted through iteration to match autopsy data on adenoma prevalence and SEER data on CRC incidence**

5

SLIDE 5 NOTES: Many assumptions went into building the model. The most important are listed in this slide. First, we assumed that all colorectal cancers originate from pre-existing adenomas. Recently we have talked to many experts in the field, and for the most part this idea is widely accepted. The real question is whether all pre-cancerous adenomas can be seen on optical colonoscopy. If not, then some cancers would appear to emerge de-novo, when in fact they have not been detected in their pre-cancerous state.

Our models assume that all adenomas progress to colorectal cancer, but some progress relatively quickly while others progress very slowly. The alternative hypothesis – that some never progress to cancer –is not inconsistent with our model, which assigns a such a long transition time to slow-growing polyps that they are very unlikely to progress in a person's lifetime.

The dwelling times for these adenomas (the slow-growing ones) were fitted through an iterative process to match the autopsy data on adenoma prevalence and SEER on SEER data on CRC incidence. The fit resulted in a mean progression time for fast-progressing adenomas of 26 years; and for slow-progressing adenomas of 75 years.

The rate of progression from asymptomatic to symptomatic CRC is based on a previous mathematical analysis using the prevalence of malignant polyps to predict the mean latency period before CRC becomes symptomatic (4.8 years). Using this information and data on the relative prevalence of different stages of CRC at symptomatic diagnosis, distribution of times for the progression of CRC from local to regional, from local to distant, and from asymptomatic to symptomatic disease were fitted. CRC can be "discovered" within the model based on either the appearance of

symptoms or detection using a diagnostic test (e.g. colonoscopy). Stage-dependent survival following CRC diagnosis was taken directly from SEER data. Data for model calibration came from a variety of sources (Arminski and McLean, 1964; Blatt, 1961; Eide, 1986; Eide and Stalsberg, 1978; Hofstad et al., 1996; Hardcastle et al., 1996; Koretz, 1993; Kronborg et al., 1996; Mandel et al., 1993; Williams, et al., 1982; Vatn, 1982; Winawer et al., 1993).

SLIDE 6

Validation

- Content validity was provided by the thoroughness of the literature search
- Construct validity was provided by ability to fit the adenoma and CRC data simultaneously
- Criterion validity was provided by a reconstruction of the National Polyp Study and correct prediction of its outcome within acceptable error range

6

SLIDE 6 NOTES: The National Polyp Study followed a cohort of 1400 people with identified adenomas on screening for 6 years following polypectomy. They found 5 asymptomatic cancers during follow-up (Winawer et al., 1993). Our model programmed with their protocol predicted that 3+3 asymptomatic cancers would be found.
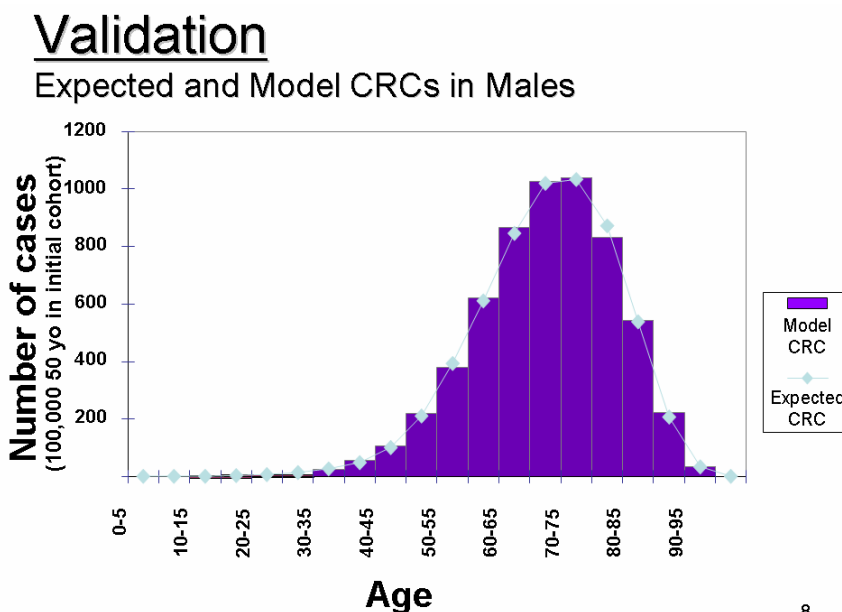
SLIDE 7

Validation
Adenoma Prevalence in Males
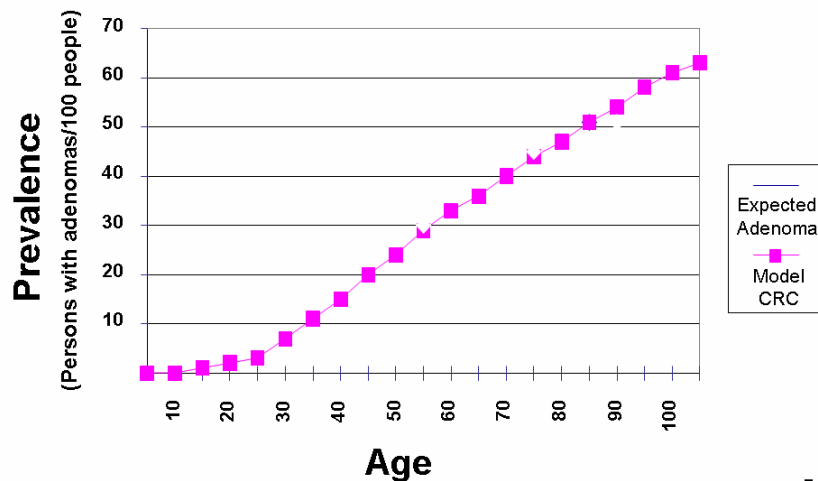
SLIDE 7 NOTES: This chart shows the fit of the model adenoma prevalence with age to the autopsy data (Arminski and McLean, 1964; Blatt, 1961; Correa et al., 1977; Eide 1986; Eide and Stalsberg, 1978; Vatn and Stalsberg, 1982; Williams et al., 1982). This graph plots prevalence in terms of persons with adenomas/100 people against patient age. We fit the data separately for males and females. This chart displays the fit for males.

SLIDE 8

Validation
Expected and Model CRCs in Males

SLIDE 8 NOTES: This chart shows (via the line) the expected number of polyps in different age ranges for a simulated population of 100,000 patients based on SEER data for 1996 (SEER, 1996).

The bars are the number of polyps produced by the model in each of these age ranges. We calculated a total chi-square of 6.04 for the fit of 12 different 5-year age ranges between 30 and 95 years of age. The goodness-of-fit of the two distributions over 12 5-year blocks between age 30 and 95 was high (Chi-Square statistic = 6.04).

We wish to point out that our model not only generates adenomas, but it also adds a background rate of non-adenomatous polyps (e.g., hyperplastic). The well-known existence of such polyps, whose prevalence has been estimated from screening colonoscopy studies, could be found during endoscopic or radiological screenings. The removal of such lesions alters costs and effectiveness of screening in two ways. First, the cost of removal and biopsy would be incurred. But, if such lesions are discovered on sigmoidoscopy or radiology, they might generate a full optical colonoscopy. That follow-up procedure could well find an adenoma by random happenstance.

SLIDE 9

Test Performance Assumptions

- Sensitivity and specificity assumptions drawn from published literature and recent reports
- The model generates a background rate of non-adenomatous polyps distributed by colon segment based on screening colonoscopy reports
- Specificity interpreted to imply specificity for polyps in the case of endoscopic or radiologic screening modalities

9

SLIDE 9 NOTES: Our assumptions about follow up and surveillance are not shown here, but I want to point out that we assume that all polyps found on sigmoidoscopy are assumed to be referred for follow-up colonoscopy without biopsy. That assumption makes sigmoidoscopy a more costly strategy than it would be if only those polyps found to meet certain high-risk criteria were included.

   In determining who is referred to followup for a positive test, and what happens to them on the followup examination, we interpret test specificity as reported in the literature to imply specificity for all polyps (adenomatous and non-adenomatous) in the case of endoscopic or radiologic screening modalities, but not in the case of FOBT tests (Allison et al., 1996; Allison et al., 1990). If a test is positive, it is considered a true positive if a polyp or cancer of any kind if found. If the polyp turns out to be hyperplastic, we do not consider the test to be a false-positive. This approach to the interpretation of specificity, which appears to differ from the interpretation of other models described at this workshop, has important implications for the outcome of cost-effectiveness analyses. We discuss this issue further later in the presentation.

SLIDE 10

Follow up and Surveillance Assumptions

- See discussion…

10

SLIDE 10 NOTES: We based our adherence (compliance) assumptions on the Behavioral Risk Factor Surveillance System. In that survey, 70 percent of people report ever having had any CRC screening modality in their lifetime. In the population, of course, people can choose one or more screening modalities from the milieu that are available. Although people report different rates of adherence for different modalities, we do not know how they would behave if they were offered just one. In the modeling context, we must evaluate a specific screening strategy that excludes the full range of choice currently available in the community.

Among those patients adherent with any screening, (i.e., 70 percent) there were different sub-populations adherent at different maximum testing frequencies. This maximum testing frequency was programmed as a patient-specific characteristic. In practice, this created rates of effective adherence that varied between testing modalities and screening strategies (e.g. effective adherence for annual FOBT = 35 percent).

SLIDE 11

Adherence Assumptions

- ~ 70% of patients report some screening test at some point in their life
- Individual capability to adhere to screening limited by frequency of screening
- Maximal adherence rate a characteristic of both the patient and the testing modality employed

11

SLIDE 11 NOTES: Patient compliance with post-polypectomy colonoscopic surveillance was set at 80 percent of the indicated population (Schoen et al., 2002). Simulated patients assigned to limited adherence with screening were more likely to be modeled as non-compliant with surveillance. Patients were assumed to be 100% adherent with surveillance after CRC.

SLIDE 12

Cost Assumptions

- Procedure costs derived form Medicare reimbursement
- Cancer treatment costs derived from published data
- Only direct health care costs associated with CRC prevention/diagnosis/treatment employed (indirect costs not included)
- Costs discounted 3% per year

12

SLIDE 12 NOTE: Procedure costs were based on Medicare reimbursement, while cancer treatment costs were derived from the experience of an HMO (Sonnenberg, et al., 2000; Taplin et al., 1995). Those costs are reported by in three phases – initial care, continuing care, and terminal care. For those patients who lived a short period of time after diagnosis, we assigned them randomly to initial care or terminal care.

SLIDE 13

Model Strengths and Weaknesses

- Strengths
  - Can incorporate quality of life adjustments
  - Can assign any number of specific attributes to each simulated patient
  - Can perform multiple interventions on each simulated patient
  - Can model the prognostic capabilities of diagnostic testing modalities
- Weaknesses
  - Programming language outdated and does not integrate well with other software packages

13

SLIDE 13 NOTES: Because of its structure as a discrete-event simulation, we can – and in other contexts have – include quality of life adjustments (Ness et al., 2000). In fact, we can assign any number of specific attributes (e.g., sex, race, age, family history) to each simulated patient. The model can accommodate multiple interventions across the full range of prevention, screening and treatment. One could examine, for example, the overall cost-effectiveness of decrease cancer risk at the same time as screening individuals for cancer.

We can also model the prognostic capabilities of diagnostic testing modalities. Since each individual is assigned a unique risk of developing colorectal cancer, a patient with higher risk of developing colorectal cancer is likely to have more or larger colorectal adenomas, which then allows us to determine whether or not he or she should enter surveillance. That way, patients who end up in surveillance are usually a higher-risk group than people who are not.

SLIDE 14

Addendum

The Effect of Ignoring Non-adenomatous Polyps in Modeling CRC Screening

14

SLIDE 14 NOTES: The following remarks relate to issues we discovered when we first received the results of the pre-workshop modeling exercise. (Note: Michael Pignone summarizes those results in another presentation.)

We noticed that, when all parameters were standardized to values provided by the organizers of the pre-workshop exercise (Run #6-see Pignone presentation), the Vanderbilt model's lifetime costs appear to be much higher than the lifetime costs reported by all other models.

In piecing together the reasons for this seeming anomaly, we realized that our model explicitly recognizes that there exists a non-negligible prevalence of non-adenomatous polyps. Some screening tests detect such lesions without being able to differentiate between them and adenomas. In particular, radiology and sigmoidoscopy would identify them as polyps requiring follow up. When that is the case, recognizing their existence in a model generates additional follow-up procedures and additional costs. In addition, there may be changes in the effect on years of life lived for reasons described below.

SLIDE 15

The Approach

- We repeated the exercise for Run # 6 (all assumptions standardized).
- In the repeat, we assumed no non-adenomatous polyps.
- We compared the results with those achieved in the first comparisons, as described by Michael Pignone.
- We present the results of the new analysis beside those of the first comparisons.
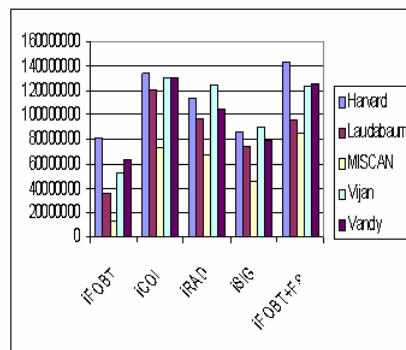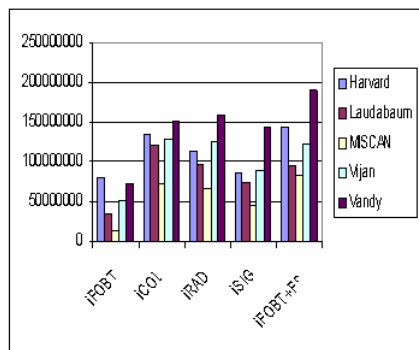
15

SLIDE 15 NOTES: In the "new" runs, we assumed that all polyps are adenomas and that their prevalence and incidence are given by the autopsy studies. Thus, we eliminated non-adenomatous polyps from the model.

SLIDE 16

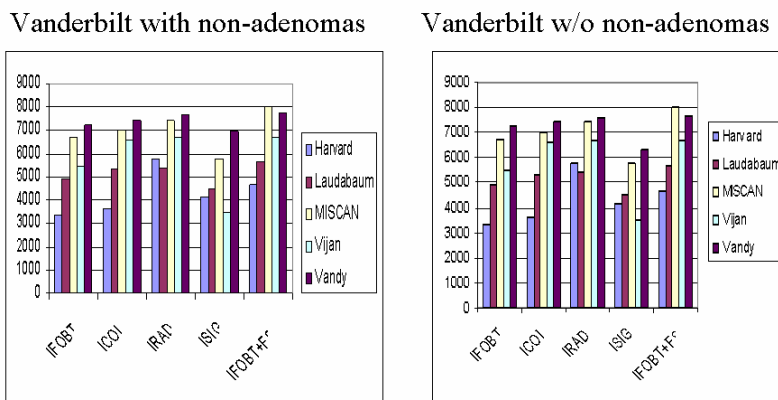Lifetime costs
with/ and without non-adenomatous polyps

SLIDE 16 NOTES: The left-hand chart shows the comparisons among models when Vanderbilt's model assumed that non-adenomatous exist and will be detected by certain screening tests. The right-hand chart shows the results when the Vanderbilt model excludes non-adenomas. When we removed the non-adenomatous polyps from the model, the costs generated by the Vanderbilt model, especially for the strategies involving sigmoidoscopy were more in line with the costs recognized for the rest of the group.

SLIDE 17 (BLANK)

SLIDE 17 is blank.

SLIDE 18

SLIDE 18 NOTES: One would expect that eliminating non-adenomatous polyps from the model could affect years of life lived, first, because some individuals undergoing sigmoidoscopy or radiology would not be sent for followup colonoscopy. But, followup colonoscopy can detect adenomas serendipitously in the proximal colon. On the other hand, there would be a small increase in effectiveness because of a decrease in mortality from complications of followup colonoscopy. The attached charts show that there is, on balance, decreased effectiveness associated with the strategies that involve sigmoidoscopy when non-adenomatous polyps are excluded from the model.

SLIDE 19

Comparison of Outcomes: Excluding vs Including Non-Adenomatous Polyps

Strategy	% Chg in Incremental Cost	% Chg in Incremental L-Y
FS q5	-45.07	-9.04
FOBT q1	-14.03	0
RAD q5	-38.00	-1.16
COL q10	-14.18	0
FOBTq1/FSq5	-34.16	-1.07

18

SLIDE 19 NOTES: The chart in this slide shows that excluding non-adenomas has the largest incremental effect on lifetime costs for screening tests that identify lesions by their physical structure (radiology and sigmoidoscopy) and require followup colonoscopy for confirmation and removal. Still, the FOBT and colonoscopy tests were affected somewhat by removal of the non-adenomatous polyps from the model. That is because any positive FOBT would lead to a followup colonoscopy, where non-adenomatous lesions would be discovered serendipitously and removed. With colonoscopy, those lesions would be found during the screening test and removed

The small reductions in effectiveness from excluding non-adenomatous polyps are also concentrated in the strategies involving radiology or flexible sigmoidoscopy.

To conclude, the change in effectiveness is dwarfed by the change in cost that results from excluding non-adenomas from the universe of colorectal lesions in our model.

SLIDE 20

ICERs All Standardized

	Harvard	Laud	Miscan	Vijan	Vandy (old)	Vandy (new)
FSq5	SD	SD	SD	SD	SD	SD
FOBTq1	$11,632	$7,272	$5,980	$9,676	$10,073	$8,659
RADq5	WD	SD	WD	SD	$192,623	$209,906
COLq10	SD	SD	SD	SD	WD	SD
FOBTq1 +FSq5	$99,977	$79,920	$55,878	$56,969	$426,956	$355,347

19

SLIDE 21 NOTES: The present slide shows that the incremental cost-effectiveness ratios changed, but the undominated strategies stayed the same when the non-adenomatous polyps were removed.

REFERENCES

Allison JE, Feldman R, Tekawa IS. 1990. Hemoccult screening in detecting colorectal neoplasm: sensitivity, specificity, and predictive value. Long-term follow- up in a large group practice setting. *Ann Intern Med*. 112: 328–333.

Allison JE, Tekawa IS, Ransom LJ, Adrain AL. 1996. A comparison of fecal occult-blood tests for colorectal-cancer screening. *N Engl J Med*. 334: 155–159.

Arminski TC, McLean DW. 1964. Incidence and distribution of adenomatous polyps of the colon and rectum based on 1,000 autopsy examinations. *Dis Colon Rectum*. 7: 249–261.

Blatt LJ. 1961. Polyps of the colon and rectum: incidence and distribution. *Dis Colon Rectum*. 4: 277–282.

Correa P, Strong JP, Reif A, Johnson WD. 1977. The epidemiology of colorectal polyps: prevalence in New Orleans and international comparisons. *Cancer*. 39: 2258–2264.

Behavioral risk factor surveillance system online prevalence data, 1995-2001. http://apps.nccd.cdc.gov/brfss; 5/5/2003.

Eide TJ, Stalsberg H. 1978. Polyps of the large intestine in Northern Norway. *Cancer*. 42: 2839–2848.

Eide TJ. 1986. The age-, sex-, and site-specific occurrence of adenomas and carcinomas of the large intestine within a defined population. *Scand J Gastroenterol*. 21: 1083-1088.

Hardcastle JD, Chamberlain JO, Robinson MH, Moss SM, Amar SS, Balfour TW et al. 1996. Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. *Lancet*. 348:1472–1477.

Hofstad B, Vatn MH, Andersen SN, Huitfeldt HS, Rognum T, Larsen S, Osnes M. 1996. Growth of colorectal polyps: redetection and evaluation of unresected polyps for a period of three years. *Gut*. 132: 449–456.

Koretz RL. 1993. Malignant polyps: are they sheep in wolves' clothing? *Ann Intern Med*. 118: 63–68.

Kronborg O, Fenger C, Olsen J, Jorgensen OD, Sondergaard O. 1996. Randomised study of screening for colorectal cancer with faecal- occult-blood test. *Lancet*. 348: 1467–1471.

Mandel JS, Bond JH, Church TR, Snover DC, Bradley GM, Schuman LM et al. 1993. Reducing mortality from colorectal cancer by screening for fecal occult blood. Minnesota Colon Cancer Control Study. *N Engl J Med*. 328: 1365–1371.

Ness RM, Klein R, Holmes AM, Dittus R. 2000. At what age is one-time colonoscopic screening for colorectal cancer most cost-effective? *Am J Gastroenterol*. 95: 1800–1811.

Schoen RE, Wissfeld JL, Trauth JM, Ling BS, Hayran M. 2002. A population-based, community estimate of total colon examination: the impact of compliance with screening for colorectal cancer. *Am J Gastroenterol*. 97: 446–451.

SEER cancer incidence public-use database, 1973-93. 1996. Department of Health and Human Services: Bethesda, MD.

Sonnenberg A, Delco F, Inadomi JM. 2000. Cost-effectiveness of colonoscopy in screening for colorectal cancer. *Ann Intern Med*. 133: 573–584.

Taplin SH, Barlow W, Urban N, Mandelson DJ, Ichikawa L, Nefcy P. 1995. Stage, age, comorbidity, and direct costs of colon, prostate, and breast cancer care. *J Natl Cancer Inst*. 87: 417–426.

Vatn MH, Stalsberg H. 1982. The prevalence of polyps of the large intestine in Oslo: an autopsy study. *Cancer*. 49: 819–825.

Williams AR, Balasooriya BAW, Day DW. 1982. Polyps and cancer of the large bowel: a necropsy study in Liverpool. *Gut*. 23: 835–842.

Winawer SJ, Zauber AG, Ho MN, O'Brien MJ, Gottlieb LS, Sternberg SS et al. 1993. Prevention of colorectal cancer by colonoscopic polypectomy. The National Polyp Study Workgroup. *N Engl J Med*. 329: 1977–1981.

Appendix H

Overview of the Vijan
Colorectal Cancer Screening Model
Sandeep Vijan, M.D., M.S.

SLIDE 1

Appendix H
Overview of the Vijan
Colorectal Cancer Screening Model

Sandeep Vijan, MD, MS
Assistant Professor of Internal Medicine
University of Michigan

Institute of Medicine, January 26-27, 2004

1

SLIDE 1 NOTES: We used this model in a paper published in 2001 (Vijan et al., 2001), but since that time we have made improvements which will be described here.

I would like to acknowledge Erica Wong, who assisted me on early stages of this project and Rodney Hayward and Tim Hofer, who provided guidance along the way.
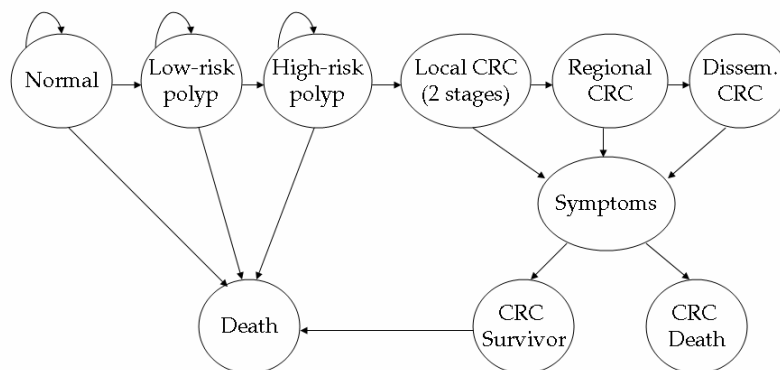
SLIDE 2

Model Structure

- Markov cohort simulation
  - Endpoints
    - Total cost, life expectancy
    - Colorectal cancer incidence, mortality
  - Designed to simulate natural history of CRC
  - Currently based on US overall rates

2

SLIDE 2 NOTES: Our model uses a Markov cohort structure that produces expected endpoints of total lifetime costs, life expectancy, and colorectal cancer incidence and mortality. It is designed to simulate the natural history of colorectal cancer based on the experience of the U.S. population. We did not stratify the analysis by sex or race.

SLIDE 3

Model Structure

SLIDE 3 NOTES: The graph shown here shows a simplified view of how we modeled the natural history of CRC in the absence of screening. Patients are distributed initially into cohorts based on prevalence data from the literature. A percentage of the population is assigned to each of the different states at age 50. In the absence of screening, all CRCs are detected through the presentation of signs and symptoms. What happens after detection – whether the patient dies from CRC or from some other cause – depends on the stage at which the symptoms developed.

SLIDE 4

Natural history assumptions

- Polyps
  - Prevalence based on autopsy studies
  - Polyp to cancer dwell time = 10 years
- Cancer
  - CRC incidence based on SEER data from early 1990s
  - Initially assumed 75% of CRC from polyps; updated model assumes 100%
  - 2 years in local cancer, 1 in regional
- Overall mortality rates also based on older data to match SEER dates
  - May account for lower total LE in our model

4

SLIDE NOTES 4: As with other models presented today, the prevalence of polyps is based on autopsy studies. We assumed a fixed time of transition from polyp to cancer of 10 years, based on the sigmoidoscopy studies that have shown a 10-year degree of protection from a sigmoidoscopy.

Cancer incidence in this model was based on SEER data from the early 1990s. We chose that period because data from earlier decades might not reflect changes in risk factors and natural history over time. However, incidence data from the 1990's might be affected by the more frequent use of CRC screening in the 1990's. In our published work we assumed that 75 percent of cancers arise from polyps that are visible on colonoscopy. The rest arise either de novo or through flat adenomas which are not seen. More recently, however, we have changed our assumptions to 100 percent arising from polyps. That change has a major impact on the relative cost-effectiveness of different strategies, because strategies that are better at detecting cancers than at detecting pre-cancerous polyps (e.g., FOBT) do not compare as favorably with strategies that can detect pre-cancerous polyps.

We assume a fixed dwelling time for a localized CRC of 2 years, and for a regional lesion of 1 year. We also assume that all individuals with disseminated cancer are symptomatic in the year they transition to that stage.

Our mortality rates are taken from life tables for the same time period as the incidence data. Our life-expectancy for the population is therefore lower, in the absence of screening than other models. Over the past 10 years, life expectancy in the US has

increased by about 2 years in a cohort of 50-year-old individuals. That 2-year difference leads to substantially lower life-expectancies in our model.

SLIDE 5

Natural history assumptions

- Calculated transition probabilities
  - Polyp incidence: calculated to match observed age-specific polyp prevalence
  - Polyp to cancer conversion rate: calculated to match observed age-specific cancer incidence
  - Symptomatic presentation of cancer to match SEER data on stage at presentation

5

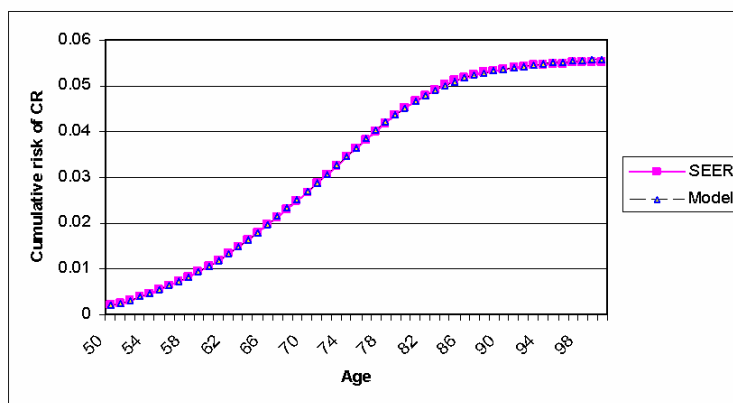SLIDE 5 NOTES: No notes.

SLIDE 6

Polyp prevalence
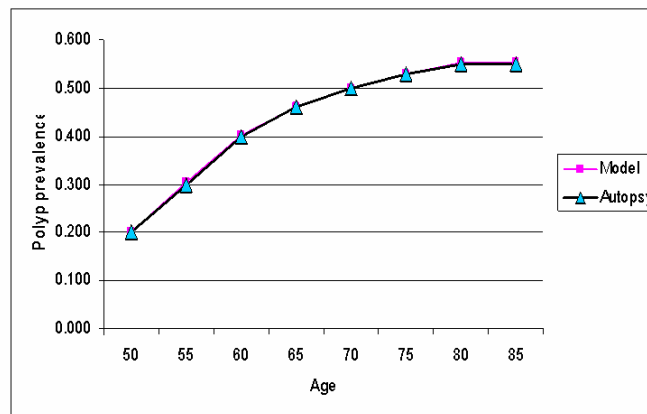
6

SLIDE 6 NOTES: The chart in this slide compares the model predictions of polyp prevalence with the autopsy data. Although we calibrated our model using visual inspection, the resulting close fit between observed and predicted values suggests a high degree of calibration.

SLIDE 7

Cumulative CRC risk

7

SLIDE 7 NOTES: This chart compares our model prediction for cumulative CRC risk in the population with the SEER cancer incidence data from age 50 through 100. The close calibration for CRC as a whole is also seen for each stage of CRC (data not shown).

SLIDE 8

Assumptions: Compliance

- Original model: base-case fixed across tests
- Updated model: follow-up colonoscopy independent of initial screening test
- FS and FOBT-based strategies negatively impacted if compliance with follow-up colonoscopy < 100%
  - If fixed, at 25% adherence, effectiveness reduced by ~94%
  - 25% adherence with baseline test, f/u colonoscopy 75%, effectiveness reduced by ~81%

8

SLIDE 8 NOTES: We model compliance as an all-or-nothing phenomenon. That is, a person either adheres perfectly to the screening strategy or does not participate at all. In our most recent version of the model (used for the exercises in this workshop) we assumed that follow-up colonoscopy after a positive screening test is independent of the initial screening test. At present, we believe the best estimate of follow-up compliance is 75%.

We have found that that there are major impacts on the cost-effectiveness of strategies involving sigmoidoscopy and FOBT if adherence to follow-up is assumed to be less than 100 percent. For example, as shown above, assuming 75% compliance with follow-up among individuals who comply with the screening test reduces the effectiveness of FS and FOBT screening by 81 percent compared with perfect compliance. That is less than the reduction in effectiveness when compliance with follow-up is set at 25%.

SLIDE 9

Assumptions: follow-up

- Base case
  - Initial model assumed all adenomas go to colonoscopy and surveillance
  - Current model assumes only adenomas >=10mm go to colonoscopy and surveillance
- Model is flexible
  - Can vary need for f/u colonoscopy and surveillance for polyps based on 3 size cutoffs: <0.5 mm; 5-9 mm; >=10mm

9

SLIDE 9 NOTES: In our published study, we assumed that all adenomas discovered by sigmoidoscopy would go to follow-up by colonoscopy and ultimately to surveillance. However, in the present version we assume that only those greater than or equal to 10 mm would go on to follow-up and surveillance. That change has an enormous impact on the cost of this intervention.

SLIDE 10

Assumptions: Cost

- Costs of screening tests from Medicare reimbursement schedule (total RVUs + facility expense)
  - Exception: FOBT no longer reimbursed; previously reimbursed at 0.5 RVU (~$18)
- Discounted costs of cancer care over lifetime are accrued at diagnosis
- Can include differential costs for cancer mortality but do not in base case
- Costs (and life yrs) discounted at 3% annually

10

SLIDE 10 NOTES: We used the Medicare reimbursement schedule to estimate test costs. FOBT is not reimbursed as a physician expense, but is paid for as a laboratory test. We assumed a cost of $18 for FOBT.

SLIDE 11

Assumptions: perspective

- Third party payer perspective
- Did not include
  - Future costs of medical illness
  - Lost productivity costs for tests
    - Colonoscopy – patient + driver
    - Sigmoidoscopy – patient only
- Did not adjust for quality of life

11

SLIDE 11 NOTES: The Panel on Cost-Effectiveness in Medicine (Gold et al., 1996) recommends that cost-effectiveness analyses take a societal perspective. As with all of the models described in the current workshop, we took a third-party payer approach with respect to measuring costs. That means that we did not include the value of lost work time (productivity costs) and other costs (e.g., transportation) associated with obtaining screening, follow-up and surveillance examinations. The productivity costs associated with colonoscopy are likely to be disproportionately higher, since that test requires a separate driver to transport patients to and from the facility because the patient must be sedated for the procedure.

SLIDE 12

Validation

- Minnesota FOBT study
  - Replaced base estimates with test/population characteristics of the study
    - Study: 33% reduction in CRC mortality
    - Model: 39% reduction in CRC mortality
  - We project lower benefits for flex sig than retrospective cohorts (~50% vs. 75% mortality reductions)

12

SLIDE 12 NOTE: Our main validation exercise was to compare the results of our model against those for the Minnesota FOBT trial (Mandel et al., 1993; Mandel et al., 1999). When we applied the test and population characteristics in that study, we predicted a 39% reduction in CRC mortality over a 13-year follow-up. Our model predicted a greater decrease in the incidence of colorectal cancer than occurred in the Minnesota trial, which may account for the differences in mortality that we found (Mandel et al., 2000). However, our results fall within the 95 percent confidence interval for the Minnesota study's observed mortality reduction.

Compared with retrospective cohort studies that examined flexible sigmoidoscopy, we predicted lower mortality on the whole. We are currently examining why our model differed from those studies.

SLIDE 13

Model limitations

- Did not take a perspective consistent with standards for CEA
- Markov model limits flexibility
  - Not able to model compliance conditional on prior compliance
  - Test results are independent when they should be correlated, especially for those non-compliant with a follow-up colonoscopy
  - Fixed dwell times
  - Variation in polyp or cancer risk difficult to estimate as requires full model recalibration

13

SLIDE 13 NOTES: No notes.

REFERENCES

Gold MR, Siebel JE, Russell LB, Weinstein MC, Editors. 1996. *Cost-Effectiveness in Health and Medicine.* New York: Oxford University Press.

Mandel JS, Bond JH, Church TR, Snover DC, Bradley GM, Schuman LM, Ederer F. 1993. Reducing mortality from colorectal cancer by screening for fecal occult blood. *N Engl J Med.* 328(19): 1365–1371.

Mandel JS, Church TR, Bond JH, Ederer F, Geisser MS, Mongin SJ, Snover DC, Schuman LM. 2000. The effect of fecal occult-blood screening on the incidence of colorectal cancer. *N Engl J Med.* 343(22): 1603–1607.

Mandel JS, Church TR, Ederer F, Bond JH. 1999. Colorectal cancer mortality: Effectiveness of biennial screening for fecal occult blood. *J Natl Cancer Inst.* 91(5): 434–437.

Vijan S, Hwang EW, Hofer TP, Hayward RA. 2001. Which colon cancer screening test? A comparison of costs, effectiveness, and compliance. *Am J Med.* 111(8): 593–601.

Appendix I

Cost-Effectiveness Analyses of Colorectal Cancer Screening: Results from a Pre-conference Modeling Exercise
Michael Pignone, M.D., M.P.H.

SLIDE 1

Appendix I
Cost-Effectiveness Analyses of
Colorectal Cancer Screening: Results
from a
Pre-conference Modeling Exercise

Michael Pignone, MD, MPH

University of North Carolina- Chapel Hill
RTI-UNC Evidence-based Practice Center

1

SLIDE 1 NOTES: I would like to thank the following people for their work and advice on this exercise or on a previous review conducted for the US Preventive Services Task Force: Judy Wagner, Louise Russell, Martin Brown, Somnath Saha, Jeanne Mandelblatt, Tom Hoerger, Steve Teutsch, and all of the modelers who participated in this exercise.

118

SLIDE 2

Aims

- To compare several cost-effectiveness analyses of colorectal cancer screening.
  - Gain insight into reasons for different results
  - Determine areas for future research focus
  - Compare different screening strategies
- To use insights from this exercise to better inform future modeling and direct future CRC research efforts

2

SLIDE 2 NOTES: The aims of the pre-Workshop modeling exercise, as I see it, were two-fold: The first was to compare the several different cost-effectiveness analyses of colorectal cancer screening. Such a comparison has three motivations: to gain insight into reasons for different results; to determine areas for future research focus; and potentially to learn something about how different screening strategies stack up against one another. The last motivation, however, is less important than the first two.

The second objective is to use insights from this exercise to better inform future modeling and direct future CRC research efforts. The exercise should lead to better understanding of our parameters, better models, and an identification of questions or issues that need to be incorporated into the models.

SLIDE 3

Background

- Colorectal cancer is an important disease that appears amenable to screening
- Several different methods screening tests are available, each supported by different evidence
- Multiple high quality cost-effectiveness analyses have been performed
- Variation in results warrants further study

3

SLIDE 3 NOTES: As with some other cancers, colorectal cancer is, of course, an important disease that is amenable to screening. However, unlike many conditions amenable to preventive interventions, there are several different screening tests available, each supported by its own body of evidence on effectiveness, risks and costs. In other areas of cancer screening there is usually one dominant modality. With CRC screening, several modalities are available.

Also, unlike many other areas, there are several recent high-quality published cost-effectiveness analyses which reached different conclusions about the relative merits of alternative screening strategies. That fact provides an interesting opportunity, because it offers us a chance to explore how that variation might arise, and whether the variation is there for good reasons, or whether we should try to reduce the variation through standardization of methods and assumptions.

SLIDE 4

Systematic Review - 2000

- All 7 models found that any of the main CRC screening strategies was cost-effective compared with no screening, with C/E ratios usually under $30,000 / life-yr saved

- Models reached different results as to the most effective and cost-effective strategy

- Variation likely due to differences and uncertainties in input parameters- no easy way to sort out these factors

4

SLIDE NOTES 4: A precursor to this exercise was the work my colleagues and I did for the US Preventive Services Task Force in 2000 (Pignone et al., 2002). At that time, we reviewed seven published models. All seven found that any of the main screening strategies for colorectal cancer were cost-effective compared with no screenings. The cost-effectiveness of any screening strategy compared with doing nothing was generally below $30,000 per year of life added across all models.

The models, however, did reach some different results as to the most effective and most cost-effective strategies. Some of those results were surprising to us.

We also concluded that the variations were likely due to differences and uncertainties in input parameters, but it was impossible to sort out these factors from the published studies. We called for an exercise similar to the one undertaken for this workshop.

SLIDE 5

Methods

- **Each modeler was asked to examine the following strategies:**
  - No screening
  - FOBT annually + SIG every 5 years
  - SIG every 5 years
  - Radiological Test every 5 years
  - COL every 10 years
  - FOBT annually

5

SLIDE 5 NOTES: Here is a brief description of the methods used in the pre-workshop modeling exercise.

Each modeler was asked to analyze 5 screening strategies, as well as no screening, as listed above.

The prototype radiological screening test was defined to have characteristics somewhere in between barium enema, which is relatively inexpensive, and virtual colonoscopy, which is more sensitive but more expensive.

SLIDE 6

Methods

- Standardized inputs designated for:
  - Test and treatment costs
  - Test Performance
  - Adherence
  - Surveillance
- Each modeler was asked to make 10 "runs" of their models with different combinations of the parameters standardized

6

SLIDE 6 NOTES: We then specified standardized values for inputs in the four categories listed above. The modelers were asked to analyze each of the six strategies 10 times, with each run involving a different combination of original or standardized parameter values.

SLIDE 7

Calculation of C/E ratios

- Tabulated life-years gained from least to most effective intervention
- Strongly dominated strategies (less effective and more costly) and weakly dominated strategies (less effective and poorer C/E ratio) were eliminated
- Incremental C/E ratios calculated
- Preferred test determined at different C/E thresholds (20,40,60,80,100K / life-yr saved)

7

SLIDE 7 NOTES: For each run, we did the following: We ordered the years of life saved for every strategy from lowest to highest. We identified the strongly dominated strategies – those which were both less effective and more costly than at least one other strategy. Strongly dominated strategies were eliminated. Of the remaining strategies, we identified those that were weakly dominated -- they were both less effective and their costs per year of life added were higher than at least one other strategy. The remaining strategies constitute the undominated set.

We then calculated the incremental cost-effectiveness ratio: the incremental costs per incremental year of life added by moving from the least (undominated) strategy to the next least (undominated) strategy, and so on.

We designated a "preferred strategy" for any cost-effectiveness limit as the one with the highest effectiveness (years of life added) whose incremental cost-effectiveness ratio meets a given limit.

SLIDE 8

Basic Assumptions

- Discount Rate = 3%
- No quality adjustment
- Start screening at age 50, end at age 80
- Cohort of 100,000 average –risk adults
- Report life-years and costs to age 85
- Use 2003 dollars

8

SLIDE 8 NOTES: This and the next 5 slides review the standardized assumptions in each general area. Here are some basic assumptions that were common to all runs and all strategies.

SLIDE 9

Assumptions: Costs

- FOBT: $10
- Screening Sig: $200 ($375)
- Radiology test: $200
- Screening Colonoscopy: $625 ($900)
- Costs of treating cancer
  - Local: $24,000
  - Regional: $31,000
  - Distant: $40,000

9

SLIDE 9 NOTES: No notes.

SLIDE 10

Assumptions: Test performance

- FOBT
  - Sensitivity Cancer 40%
  - Sensitivity Polyps 10%
- Sigmoidoscopy reach: 50% of colon
- Sigmoidoscopy
  - Sensitivity Cancer: 95%
  - Sensitivity Polyps: 85%
- Colonoscopy
  - Sensitivity Cancer: 95%
  - Sensitivity Polyps: 85%
  - Specificity: 100%
- Radiology test
  - Sensitivity cancer: 80%
  - Sensitivity polyps: 70%
  - Specificity: 90%

10

SLIDE 10 NOTES: No notes.

SLIDE 11

Assumptions: Complications

- COL perforation: 0.1%
- COL mortality: 0.01%
- SIG perforation: 0.00%
- RAD perforation: 0.00%

11

SLIDE 11 NOTES: Note that we did not model more complex assumptions regarding complications, such as the possibility of bleeding (short of perforation) with colonoscopy, or other complications, such as a patient who is falls and breaks a bone after colonoscopy.

SLIDE 12

Assumptions:
Follow-up and Surveillance

- All + FOBT leads to COL
- All + SIGs lead to COL
- Positive RAD leads to COL
- Positive COL leads to polypectomy
- All adenomas have surveillance every 5 years until death or age 80

12

SLIDE 12 NOTES: All patients with positive FOBT tests would receive a follow-up colonoscopy.

All patients with positive sigmoidoscopy would receive a follow-up colonoscopy.

All patients with positive radiology test would receive a follow-up colonoscopy

All patients with a polyp found on colonoscopy screening would have the lesion removed as part of that procedure.

All patients with adenomas found on screening and removed in screening or follow-up would be entered into a surveillance program requiring a full colonoscopy every 5 years until death or until the patient reaches age 80.

SLIDE 13

Assumptions: Compliance

- Assume 100% compliance

13

SLIDE 13 NOTES: When we standardized on assumptions about compliance, we asked modelers to assume that all individuals would be fully compliant with all screening, follow-up and surveillance tests. That assumption is a poor description of reality, but it provided a level playing field for all procedures.

SLIDE 14

Standardization Analyses

	1	2	3	4	5
Costs	O	S	O	O	O
Test Performance	O	O	S	O	O
Surveillance	O	O	O	S	O
Compliance	O	O	O	O	S
Basic Assumptions	S	S	S	S	S

14

SLIDE 14 NOTE: In this and the next chart, the rows depict the different assumptions and the columns depict the specific run. "S" means that the parameters in a specific run and input group (for example, in run number 2, and the "Cost" assumptions) were set to the standardized values we specified. "O' means that the parameters in a specific run and input group (for example, in run number 3 and the "Cost" assumptions) were set to the values in the modeler's originally published or current version of the model.

Run number 1 represents the original assumptions across all four parameter areas.

SLIDE 15
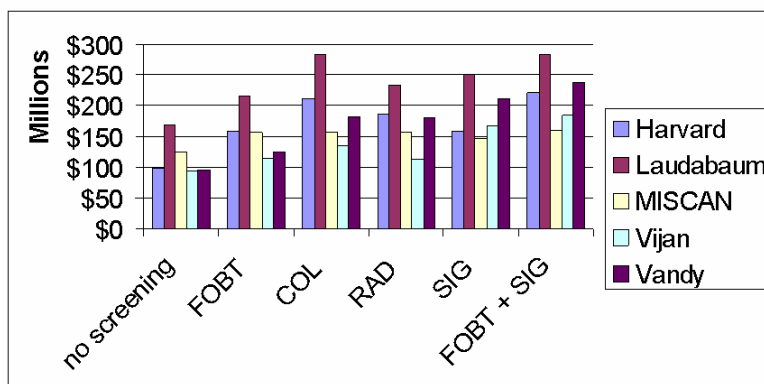
Standardization Analyses - 2

	6	7	8	9	10
Costs	S	S	S	S	O
Test Performance	S	S	S	O	S
Surveillance	S	S	O	S	S
Compliance	S	O	S	S	S
Basic Assumptions	S	S	S	S	S

15

SLIDE 15 NOTES: Run 6 standardizes across all parameter groups. We call that the fully standardized run.

SLIDE 16

Results:
Costs- Original Assumptions

16

SLIDE 16 NOTES: Now for the results of the exercise.

The current chart shows – for the original assumptions (Run 1) -- the lifetime cost in a population of 100,000 50-year old individuals of screening, follow-up, surveillance and treatment of CRC in millions of dollars. That value is shown for each of the five screening strategies and no screening.

Within each screening test, there is fairly substantial variation under the original assumptions about the costs of screening.
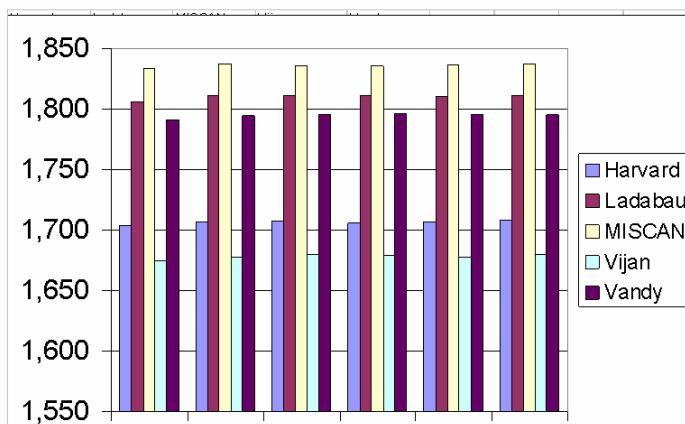
SLIDE 17

SLIDE 17 NOTES: This chart shows –for the original assumptions -- the years of life lived in a population of 100,000 50-year old individuals.

Here, too, there is a substantial variation between the different models in terms of the number of years of life that would be generated through running the model under each of the different modelers' assumptions.

SLIDE 18

Life Years c/w No Screening- Original
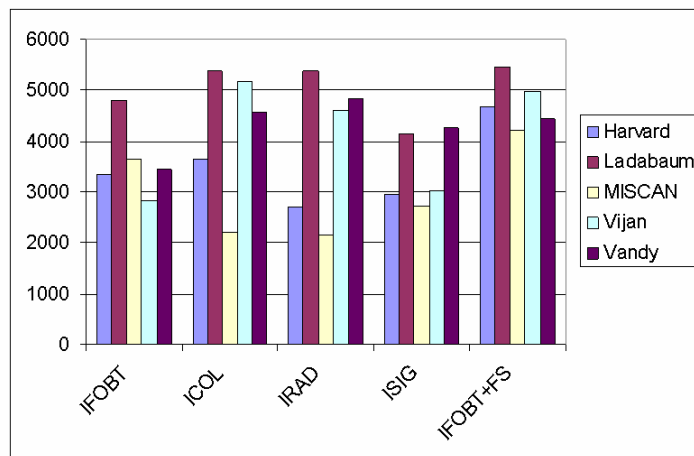
SLIDE 18 NOTES: This slide shows the years of life added, compared with the no-screening strategy, under the original assumptions (Run 1).

Although the metric has changed (from total number of years of life lived to additional years of life lived), the same pattern is replicated across the models. Substantial differences exist between models.

SLIDE 19

Costs c/w No Screening- Original
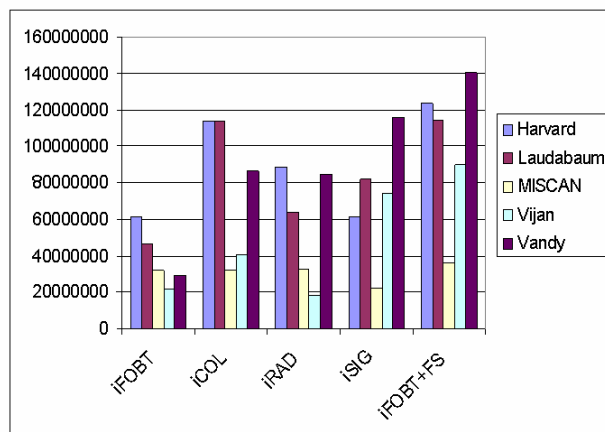
SLIDE 19 NOTES: This chart shows extra lifetime costs compared with the no-screening strategy, under the original assumptions (Run 1). Substantial differences in cost persist across models for each of the different screening strategies. a full colonoscopy every 5 years until death or until the patient reaches age 80.

SLIDE 20

Average Cost-effectiveness- Original

| Study	Harvard	Lada	MISCAN	Vijan	Vandy
FS q5	$20,681	$19,741	$8230	$24,449	$27,169
FOBT q 1	$18,347	$9,631	$8815	$7,517	$8,409
RAD q5	$20,573	$11,674	$15,054	$3,980	$18,919
COL q10	$31,211	$21,167	$14,456	$7,810	$18,940
FOBT q 1 + FS q5	$26,324	$21,000	$8,448	$18,000	$31,621

20

SLIDE 20 NOTES: Here are the average cost-effectiveness ratios under the original assumptions (Run 1). By average, I mean the ratio of additional costs to additional effectiveness when each strategy is compared with no screening.

A quick scan of these results shows that almost all of the strategies have cost-effectiveness ratios of less than $30,000, regardless of the model used. The cost-effectiveness ratios tend to vary between $10,000 and $30,000 both across screening strategies and across models.
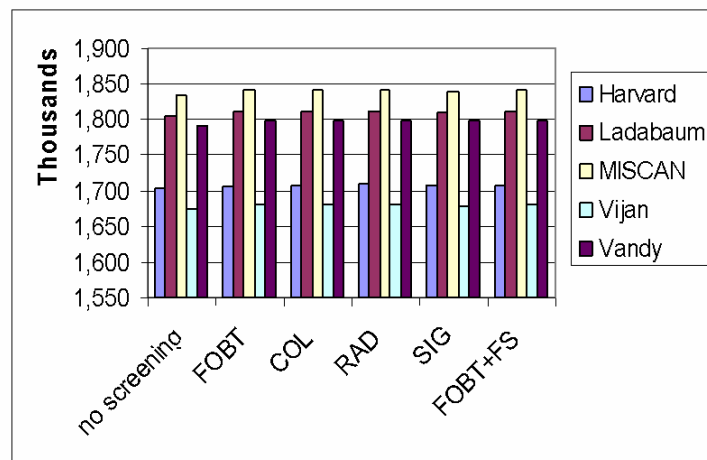
SLIDE 21

Costs – All Standardized
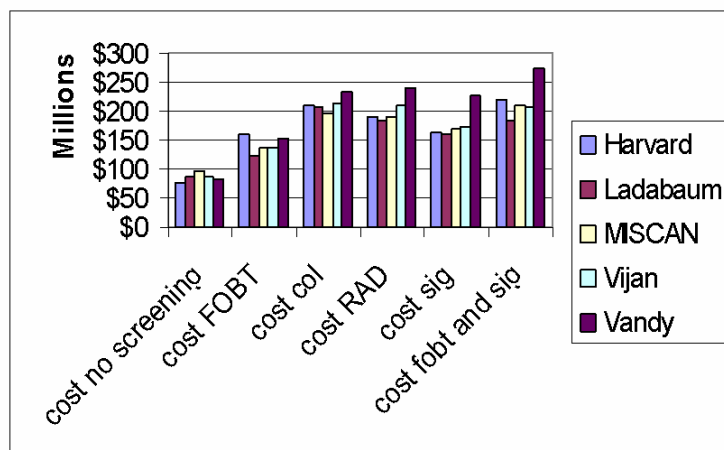
21

SLIDE 21 NOTES: This and subsequent slides provide results under the fully standardized assumptions (Run 6). All four parameter groups are standardized in this run.

The current slide shows – for the standardized assumptions (Run 6)—the lifetime cost of screening, follow-up, surveillance and treatment of CRC in millions of dollars. That value is shown for each of the five screening strategies and no screening.

Now there is much less variation across the models once assumptions have been standardized.

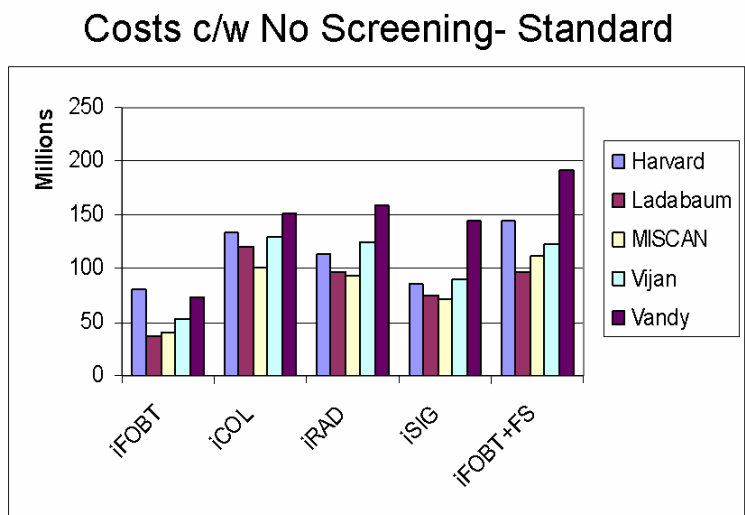**SLIDE 22**

Life-years – All Standardized

22

SLIDE 22 NOTES: This chart shows the years of life lived in a population of 100,000 50-year old individuals under the standardized assumptions.

Rough visual inspection suggests that there is probably a little less variation than there was under the original assumptions. Still, substantial variation exists in the number of years of life lived across models.

Interestingly, there appears to be more variation among the different models than there is among the different tests. These differences across models may reflect different assumptions about the natural history of CRC. Note that neither natural history nor model structure have been standardized.

SLIDE 23

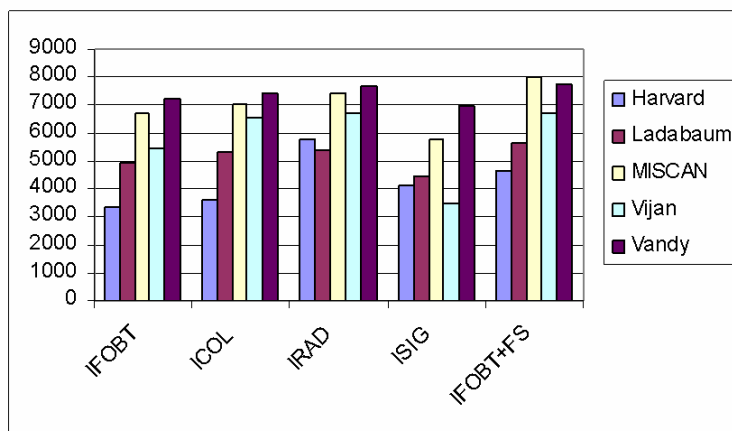Costs c/w No Screening- Standard

23

SLIDE 23 NOTES: This slide shows extra lifetime costs compared with the no-screening strategy, under the standardized assumptions (Run 6). Now there is greater similarity across models in terms of costs once inputs are standardized.

SLIDE 24

## Life Years c/w No Screening - Standard

24

SLIDE 24 NOTES: This slide shows the years of life added, compared with the no-screening strategy, under the standardized assumptions (Run 6). Variation across models is now somewhat reduced, probably because some of the differences in assumptions about natural history wash out when the metric is years of life added compared with no screening. Nevertheless, substantial differences remain across models.

SLIDE 25

## Life years c/w no screening: standardized and by model
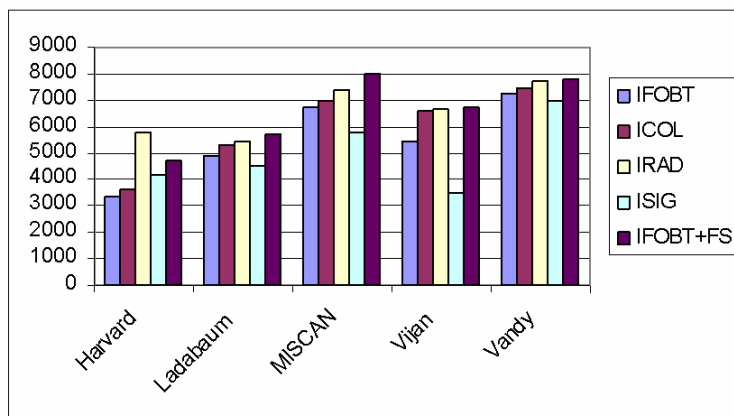
25

SLIDE 25 NOTES: In this chart and the next, the same results are grouped by model instead of by strategy. You can see that there is some variation in terms of life years saved within each model by the different strategies, suggesting that the strategies have different levels of effectiveness.

SLIDE 26

Costs c/w no screening: standardized and by model
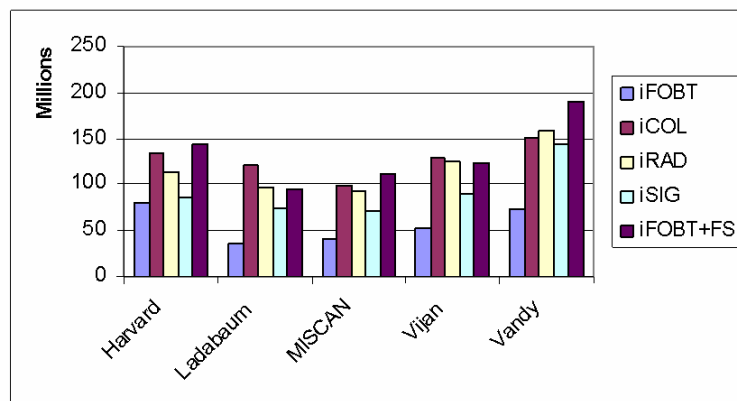
26

SLIDE 26 NOTES: This slide groups lifetime costs by model. There are now some differences across strategies for all models, but they are relatively small across different tests, with FOBT generally less costly in each model than the other screening strategies. The relative costs across strategies tend to follow pretty much the same pattern across the different models.

SLIDE 27

Average Cost-effectiveness- Standard

Study	Harvard	Lada	MISCAN	Vijan	Vandy
FS q5	$20,509	$16,689	$12,363	$25,722	$20,647
FOBT q 1	$11,632	$7,272	$5,980	$9,676	$10,073
RAD q5	$19,609	$17,953	$12,598	$18,699	$22,099
COL q10	$23,579	$22,732	$14,181	$19,695	$20,316
FOBT q 1 + FS q5	$21,381	$16,898	$13,922	$18,301	$24,570

27

SLIDE 27 NOTES: Here is the average cost-effectiveness under the standardized assumptions (Run 6). The results are quite similar to what was seen under the original assumptions for average cost-effectiveness. The results here vary from about $6,000 per life-year saved to about $25,000 per life-year saved. It appears from visual inspection that there is slightly less overall variation than we had under the original assumptions.

SLIDE 28

Incremental Cost-Effectiveness Ratios and Preferred Strategies

28

SLIDE 28 NOTES: The rest of this presentation is about incremental cost-effectiveness ratios (as opposed to average cost-effectiveness ratios) and preferred strategies.

Recall that the incremental cost-effectiveness ratio (ICER) is calculated by eliminating all strongly and weakly dominated strategies and then sorting the remaining strategies in ascending order according to years of life added compared with no screening. The incremental ratio is calculated for each strategy as the extra costs incurred per extra year of life added by moving from each strategy to the next most effective strategy.

SLIDE 29

ICERs: <u>Original</u> Assumptions

Study	Harvard	Lada	MISCAN	Vijan	Vandy
FS q5	WD	SD	$8,230	SD	SD
FOBT q 1	$18,347	$9,631	WD	SD	$8409
RAD q5	WD	$27,069	SD	$3,980	$44,936
COL q10	WD	SD	SD	$38,854	WD
FOBT q 1 + FS q5	$45,976	SD	$8,848	SD	SD

29

SLIDE 29 NOTES: Here are the ICERs under the original assumptions (Run 1). There are definite differences across models in which strategies are dominated and which are not.

For example, flexible sigmoidoscopy every five years is dominated in all but the Miscan model.

In four of the five models, colonoscopy is either weakly or strongly dominated by other tests. In the Vijan model it has a cost-effectiveness ratio of $38,000 per year of life added.

SLIDE 30

ICERs: All <u>Standardized</u>

Study	Harvard	Lada	MISCAN	Vijan	Vandy
FS q5	SD	SD	SD	SD	SD
FOBT q1	$11,632	$7,272	$5,980	$9,676	$10,073
RAD q5	WD	SD	WD	SD	$209,906
COL q10	SD	SD	SD	SD	WD
FOBT q1 + FS q5	$99,977	$79,920	$55,878	$56,969	$355,647

30

SLIDE 30 NOTES: This chart shows the ICER's under the standardized assumptions (Run 6). With assumptions standardized, the first four models get very similar results. Under the specific set of standardized assumptions made about each strategy, FOBT screening generally had ICER's between $5,000 and $12,000 per year of life saved, while flexible sigmoidoscopy every five years was dominated in all models.

Radiology and colonoscopy were dominated in most models. When they were not, they had a high incremental cost-effectiveness ratio.

Finally, annual FOBT plus flexible sigmoidoscopy under these particular standardized assumptions produced additional life years at a fairly high additional cost. The highest estimate was from the Vanderbilt model, with over $350,000 per additional life year saved. Those results merit more discussion.

SLIDE 31

Most Effective Strategy- Original

Study	$20,000	$40,000	$60,000	$80,000	$100,000
Harvard	FOBT	FOBT	FOBT/FS	FOBT/FS	FOBT/FS
Lada	FOBT	RAD	RAD	RAD	RAD
MISCAN	FOBT/FS	FOBT/FS	FOBT/FS	FOBT/FS	FOBT/FS
Vijan	RAD	COL	COL	COL	COL
Vandy	FOBT	FOBT	RAD	RAD	RAD

31

SLIDE 31 NOTES: This slide shows – for the original assumptions -- the most effective strategy (i.e., the strategy that produces the largest number of additional years of life among all non-dominated strategies) under given incremental cost-effectiveness thresholds. So, for example, the Harvard model predicts that annual FOBT is the most effective strategy among all strategies whose ICER is $20,000 or less.

There is a great deal of variation across models in which strategy is preferred at any cost-effectiveness threshold.

SLIDE 32

Most Effective Strategy- Standardized

Study	$20,000	$40,000	$60,000	$80,000	$100,000
Harvard	FOBT	FOBT	FOBT	FOBT	FOBT/FS
Lada	FOBT	FOBT	FOBT	FOBT/FS	FOBT/FS
MISCAN	FOBT	FOBT	FOBT/FS	FOBT/FS	FOBT/FS
Vijan	FOBT	FOBT	FOBT/FS	FOBT/FS	FOBT/FS
Vandy	FOBT	FOBT	FOBT	FOBT	FOBT

32

SLIDE 32 NOTES: This chart is the same as the previous chart, except that the assumptions are fully standardized (Run 6). Here, almost all of the differences across models disappear. The most effective strategy at any different threshold is the same. In fact, the only difference is the threshold level at which the FOBT with FSIG overtakes FOBT alone as the preferred strategy.

SLIDE 33

Limitations

- Limited set of parameters used- did not % cancers arising from polyps, dwell time, among others
- Only examined one set of values for standard parameters- different values may have produced different results
- Did not examine other testing strategies
- Did not account for uncertainty in estimates (i.e no sensitivity analysis)

33

SLIDE 33 NOTES: This exercise had several limitations. Some were a function of the limited time we had to design and conduct the exercise and the amount of effort that the modelers could realistically expect to make to support the exercise.

The exercise did not examine the effect of different assumptions about the natural history of colorectal cancers.

The modelers were provided with only one set of standard values for assumptions. Those standard values were not all realistic; many were selected because they would require the least amount of model redesign. True values of such parameters might be quite different, and another exercise would be warranted for such parameters.

The effects of standardizing assumptions might differ with other sets of screening strategies. It might be useful, for example, to do an exercise that includes more complex screening strategies such as one that begins with one screening test and transitions over time to another as individuals age.

Finally, there was no accounting for uncertainty in estimates. There is a method of doing sensitivity analysis, and we did not do Monte Carlo simulations to generate confidence intervals around some of our parameters. So we are dealing with point estimates here.

SLIDE 34

Conclusions

- All models found all strategies to be cost-effective compared with no screening
- Original model results showed substantial variation in ICERs and preferred strategies
- Adjustment for differences in costs, test performance, surveillance, and compliance mitigated many of the differences
- No single factor accounted for all differences, although cost adjustment was the single strongest

34

SLIDE 34 NOTES: In this chart and the next, the same results are grouped by model instead of by strategy. You can see that there is some variation in terms of life years saved within each model by the different strategies, suggesting that the strategies have different levels of effectiveness.

SLIDE 35

Implications

- Need to establish standard cost inputs
- Additional research and modeling on compliance / adherence
- Application to other health questions for which multiple models are available (e.g. treatment of hyperlipidemia)

35

SLIDE 35 NOTES: Here are some preliminary thoughts about implications of this exercise. First, it would certainly be a good idea to establish some standard cost inputs, to eliminate this major source of variation across models.

We also need additional research in modeling compliance. I believe we are still missing some of the key input parameters that would help us to more effectively and accurately model what actually happens in terms of compliance.

Finally, I would like to see this process be applied in a number of different health areas. This kind of exercise can teach us a great deal, and I hope that this can be a model for future meetings or group projects. Not only can we understand better why results differ, but we may also advance the field of modeling itself.

REFERENCES

Ness RM, Holmes A, Klein R, Greene J, Dittus R. 1998. Outcome states of colorectal cancer: Identification and description using patient focus groups. *Am J Gastroenterol.* 93(9): 1491–1497.

Pignone M, Saha S, Hoerger T, Mandelblatt J. 2002. Cost-effectiveness analyses of colorectal cancer screening: a systematic review for the U.S. Preventive Services Task Force. *Ann Intern Med.* 137(2): 96–104.

Appendix J

Recent Findings on Test Performance
Brian P. Mulhall, M.D., M.P.H.

SLIDE 1

Appendix J
Recent Findings on Test Performance

Brian P. Mulhall, MD MPH
Walter Reed Army Medical Center

1

SLIDE 1 NOTES: I was asked to cover recent evidence on test performance for all the major colorectal cancer screening tests.

I tried to take a meta-analytic approach to every test I examined, but was frustrated by the quality, quantity and heterogeneity of the various studies examining each individual test. So, my review will be largely descriptive, followed by a summary of what, in my opinion, we know about the various tests.

SLIDE 2
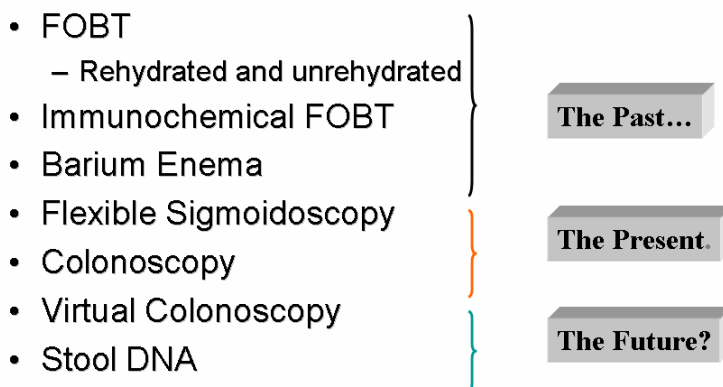
SLIDE 2 NOTES: Here is a list of what I consider tests that are viable for colorectal cancer screening.

They begin with FOBTs and barium enema. These are to some extent viewed by both clinicians and the public as part of the past.

Flexible sigmoidoscopy and, more importantly, colonoscopy, are where clinicians and the public currently see the state-of-the-art for colorectal cancer screening.

Virtual colonoscopy and Stool DNA tests are largely in the future, from an operational point of view, though they have recently received a good amount of attention in the press.

SLIDE 3

Fecal Occult Blood Tests

- FOBTs
  - GUAIAC-based
    - Hemoccult II
    - Hemoccult SENSA
    - Fecatwin S
    - Coloscreen III/VPI
    - Hemoccult Wipe
  - Immunochemical tests
    - FECA-EIA
    - Monohaem
    - HemeSelect
  - HemoQuant

Sensitivity	Specificity	PPV	NPV
60%	83	33	94
78	73	27	96
60	67	21	92
56	41	11	88
56	75	23	93
89	62	23	98
70	59	20	93
33	78	18	89
50	81	29	91

n=81

Source: Gopalswamy, 1994

3

SLIDE 3 NOTES: Fecal Occult Blood Tests (FOBTs) have been in use for several decades and provide some of the best data supporting the indication for CRC screening in order to prevent colorectal cancer (Ederer et al., 1997; Mandel et al., 1993). There are over a half-dozen different tests that are GUAIAC-based and even more that have been developed using immunochemical assays to address some of the limitations of GUAIAC-based stool studies Fecal occult blood tests can also be immunological (e.g., HemoQuant).

Ideally, we would want to see a study that compares all of these various types of tests head-to-head. In the mid-1990's Dr. Gopalswamy and colleagues attempted to do just that. (See this slide.) Unfortunately, that study suffers from a too-small sample size as well as uncertainty about the definition of the outcome measures used.

Consequently, we are left with a variety of studies of individual tests and must try to compare their outcomes.

SLIDE 4

Test Performance
FOBTs

- GUAIAC-based
  - Hemoccult II
  - Hemoccult SENSA
  - Fecatwin S
  - Coloscreen III/VPI
  - Hemoccult Wipe
  - Hemofec
  - Hemopreuve
- Unrehydrated/Re-hydrated

Sens.	Spec.	PPV	NPV
2-80%	19-97%	27-33%	94%
4-78%	47-84%	27%	96%
56%	41%	11%	88%
56%	75%	23%	93%
60%	67%	21%	92%
91%	73%	46%	97%
94%	67%	42%	98%

4

SLIDE NOTES 4: FOBTs - There is a range of results reported for the GUAIAC-based stool studies, with sensitivity for detection of polyps ranging from 5-98% and sensitivity for detection of cancer of 18-92 percent. Specificity for cancer and/or polyps ranges from 90-100 percent, but large cohort studies emphasize that a significant number of CRCs are missed by FOBTs (Bouvier et al., 2001; Lieberman et al., 2001). Results differ based on the design of the study, number of tests used, method of follow-up and the cut-off used for positive results, for example, all adenomas, adenomas > 1 cm or actual colorectal cancer (Brevinge et al., 1997; Favennec, 1992; Gopalswamy et al., 1994; Greenberg et al., 2000; Hope et al., 1996; Kewenter et al., 1988; Ko et al., 2003; Niv et al., 1995; Rockey et al., 1998; Rosen et all., 1997; Sung et al., 2003).

   In general, the overall sensitivity of GUAIAC-based FOBTs is low and they are frequently positive as a result of upper gastrointestinal lesions (Lieberman et al., 2001; Rockey et al., 1998). There is no clearly superior test among the FOBT alternatives, but rehydration of FOBTs does appear to improve sensitivity in several studies but decreases the specificity (Kewenter et al., 1988; Walter et al., 1991).

SLIDE 5

Test Performance
Immunochemical FOBTs

- **IFOBTs**

	Sens-Polyps	Sens.-Ca	PPV	NPV
– FECA-EIA	89%	62%	23%	98%
– Hemolex	70%	82%	92%	91%
– Monohaem	30-70%	44-89%	20-74%	93-97%
– HemeSelect	4-65%	49-97%	18-66%	89-96%
– HemoQuant	50%	18-81%	29%	91%
– Hemoblot	38-54%	81-94%		
– BM-test Colon	16%	20-30%	12-29%	86-97%
– Immudia-HemSp	25-54%	93%		
– OC-hemodia	4-54%	77-94%	5-18%	
– Immunohemostick	13-89%	79-98%		
– RPHA	44-70%	69-91%		
– Fecal calprotectin	37-68%	63-90%	4.5-23%	76-99%
– Magstream 1000	13-29%	57-100%	14-53%	
– FlexSure OBT	8-36%	86-99%	29-42%	86-99%

2

SLIDE 5 NOTES: IFOBTS -- There is a broad range of results for Immunochemical FOBTs (IFOBTs) impacted by the number of studies used, the population screened, the 'gold standard' used and the lesion (and size) used to define a positive result (Gopalswamy et al., 1994; Greenberg et al., 2000; Ito et al., 1996; Jeanson et al., 1994; Ko et al., 2003; Miyoshi et al., 2000; Nakama et al., 1997; Nakama et al., 1999; Nakama et al., 2001; Rosen et al., 1997; Rosen et al., 2000; Rosen et al., 1995; Sieg et al., 1998;Wong et al., 2003). For IFOBTs, the range of sensitivities for detection of polyps ranges from 4-98 percent and for colorectal cancer ranges from 18-100 percent. Reported specificity for CRC is generally greater than 95 percent.

Based on the review of the literature, there is also no clearly superior IFOBT at present time, but it does appear to have improved performance characteristics compared to GUAIC-based FOBTs in a number of studies. A recent study has argued against the assertion that IFOBTs are superior to FOBTs (Ko, et al., 2003), and others have also not demonstrated a conclusive advantage (Rosen et al. 1997; Rosen et al., 2000).

SLIDE 6

FOBT vs. IFOBT

- ## Large VA-based population study
  - Varied risk for CRC, 98 % male
  - IFOBT (FlexSure OBT) vs. FOBT (HemoccultSENSA)
    - 2965 tests ordered each; 1410/1369 returned
      - 48 % return rate
      - 66 % positive results referred for colonoscopy (50 % completed)
  - ## Test Performance

	PPV Adenoma	PPV Adenoma > 1 cm or Cancer
• IFOBT →	58 %	17 %
• FOBT →	59 %	30 %

Source: Ko et al (2003)   3

SLIDE 6 NOTES: A new article published by Ko and colleagues (Ko et al., 2003) looked at a population of veterans at varied risk for colorectal cancer, almost all male. The purpose of the study was to compare the overall performance of IFOBTs and FOBTs.

In this study, 3,000 of each test type were ordered, but only half of those in each group were returned by the patient. That equates to a 48 percent return rate. That rate is consistent with other studies in our population.

Of those who had a positive result on the test, only 66 percent were referred to colonoscopy, and of that number, only 50 percent actually received that colonoscopy. Dr. Ko's study showed that the overall Positive Predictive Value (PPV) of IFOBT and FOBT for either adenomas or cancers was about the same. For large adenomas or cancer, IFOBT was not as good as FOBT, but the difference was not statistically significant.

SLIDE 7

Test Performance
Barium Enema

DCBE	Polyps < 1 cm	Polyps > 1 cm	Cancer
Sensitivity	2-70%	50-90%	62-100%
Specificity	67-100%		90-98%

- Infrequently used in practice
- Limited data on impact on outcomes
- Varied studies
  - Symptomatic or high-risk populations (Sensitivity → 85-90%)
    - Work-up bias
  - Retrospective cohort studies (Sensitivity → 70-96%)
    - Biased
  - Surveillance populations Sensitivity → 30-81%;
    - NPS Sens 50% polyps > 1 cm
    - Selection bias
  - Tagged polyp study (specificity → 96-98% for polyps > 7-10 mm or CA)
    - Data for hyperplastic versus adenomatous polyps not reported

4

SLIDE 7 NOTES: Barium Enema --Though more widely used in the past, double-contrast barium enema (DCBE) has fallen out of favor with primary care providers, radiologists and gastroenterologists. The reason for this bias against BE is unclear, but may be related to published concerns about sensitivity (Rex et al., 1997).

Over 60 studies have looked at DCBE in a variety of settings. Studies that have examined symptomatic or high-risk patients, have determined that BE has sensitivity for polyps of 48-100 percent and for CRC of 62-100 percent (Glick et al., 1998; deZwart et al., 2001). Retrospective studies in cohorts of patients known to have CRC have shown that DCBEs done prior to the eventual diagnosis were 70-96 percent sensitive for CRC (Jensen et al., 1990; Johnson et al., 1983; Loose et al., 1974; Rex et al., 1997). Though no study has been done in a truly asymptomatic, average-risk screening population, the studies that have examined patients undergoing surveillance for polyps or in those with positive findings on previous studies have demonstrated a sensitivity of 30-81 percent, varying greatly depending on the size threshold used for detection of polyps (Steine et al., 1993; Saito et al., 1989; Brewster et al., 1994). Sensitivity for CRC was 80 percent in one study (Kewenter et al., 1995). There is limited data on the specificity of DCBE, but one study that evaluated tagged polyps determined the specificity for polyps > 10 mm to be 96 percent (Rex et al., 1986). Several additional studies showed the sensitivity to be 90-96 percent for polyps, depending on size thresholds, and for cancer or polyps the sensitivity was 98 percent in one study (Williams et al., 1974). In general, the gold standard in these studies was optical colonoscopy or flexible sigmoidoscopy (for left-sided lesions).

One study that compared DCBE to Flexible Sigmoidoscopy showed a clear superiority for Flexible Sigmoidoscopy (sensitivity for CRC 62 percent vs. 85 percent, whole bowel vs. rectosigmoid visualization, respectively) (Jensen et al., 1990). When used in combination, these studies had a sensitivity of 96 percent for all neoplasms > 1 cm and a specificity of 99 percent. Performance may be improved with multiple readings (Markus et al., 1990).

There are many biases inherent to the studies of DCBE for detection of polyps and cancer, most that would have biased estimates of sensitivity upwards (Glick et al., 1998). The most important deficit in the literature regarding this technology is the lack of studies on test performance in average-risk screening populations. Nonetheless, the studies cited suggest acceptable sensitivity and specificity that make the inadequate study of, and under-utilization of, this screening modality perplexing.

In summary, I would suggest that the sensitivity of the DCBE is somewhere in the 50-70 percent range for polyps, and perhaps higher than that for cancer. The specificity is 80 percent plus if you look at polyps and cancers combined.

SLIDE 8

SLIDE 8 NOTES: Flexible Sigmoidoscopy: This test has long been used for CRC screening, but has obvious limitations based on the length of insertion of the traditional Flexible Sigmoidoscopy (FlexSig) endoscope. It limits examination to the rectosigmoid, as fewer than 50 percent of FlexSigs are able to examine beyond the sigmoid colon (Ott et al., 1982; Osgard et al., 1998; Painter et al., 1999; Stewart et al., 1999). A great number of well-designed studies have examined the performance of this test, and it has a proven ability to reduce the incidence of CRC and improve mortality (Brenner et al., 2001; Newcomb et al., 2003; Walsh et al. 2003; Wherry et al., 1994). As such, data from FlexSig studies has been extrapolated to support the efficacy of Colonoscopy (Khullar et al., 1997; Rex et al., 2002).

In summary, overall sensitivity is probably 60 to 70 percent, but sensitivity for distal lesions within the reach of the FlexSig, it is in the 85 to 95 percent range, and specificity is reasonably high at 85 to 98 percent.

SLIDE 9

Test Performance
Colonoscopy

Colonoscopy	Polyps < 1 cm	Polyps > 1 cm	Cancer
Sensitivity	75-85%	79-100%	79-100%
Specificity	Uncertain		90-96%

- The 'Gold Standard' for diagnosing polyps and CRC
  - Visualizes the entire colon
    - Examination completed → 80-98%

  - Allows identification, removal of culprit lesions
    - Prevents 60-80% colorectal cancers

  - Sensitivity
    - Cancer → 79-100%
    - Polyps → 73-100% (85% polyps > 5 mm; 92% > 7 mm)
    - Miss rate for polyps > 8 mm → 0-26%; Polyps < 8 mm → 15-30%

  - Specificity
    - Assumed to be 100% --Likely in the range of 90-95%

6

SLIDE 9 NOTES: Colonoscopy -- Long considered the 'gold standard' for CRC screening, this test provides visualization of the entire colon and allows for removal of specimens for histopathologic assessment. Because it has been the intervention that has followed most of the aforementioned tests, the reduction in CRC and mortality from cancer has been attributed to colonoscopy.

Its sensitivity compared to radiographic studies and autopsy findings has been reported to range from 80-100 percent for CRC and from 70-100 percent for polyps (Brewster et al., 1994; Irvine et al., 1988; Pickhardt et al., 2003; Postic et al., 2002; Saito et al., 1989). Colonoscopy has a miss rate for colonic polyps ranges from 0 to 27 percent, but ranges according to size of the adenoma or cancer used as criteria and the endoscopic technique (Hixson et al., 1990; Hixson et al., 1991; Leinicke et al., 1977; Rex et al., 1997; Rex et al., 2000; Sheheda et al., 2002). The cecum cannot be reached in 0 to 18 percent of cases (Cirocco and Rusin, 1995; Lichtenstein et al., 1999; Nelson et al., 2002; Obrecht et al. 1984; Shumaker et al., 2002).

Because there has been no clear standard against which to compare colonoscopy, its true specificity has not been adequately described. However, using radiographic methods and segmentally unblinded colonoscopy, we can estimate that the specificity of colonoscopy is in the range of 90-98 percent. Use of chromoendoscopy may improve performance characteristics of colonoscopy (Eisen et al., 2002; Kiesslich et al., 2001; Lee et al., 2003; Tsuda et al., 2002).

To summarize, sensitivity for polyps is somewhat greater than 75 percent, and for cancer, closer to 100 percent. Specificity for polyps is unknown; for cancer it is in the 90-96 percent range.

SLIDE 10
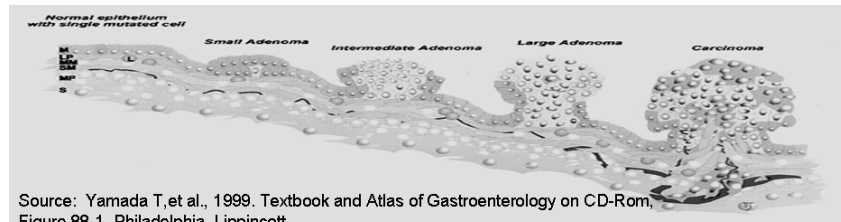
Source: Yamada T,et al., 1999. Textbook and Atlas of Gastroenterology on CD-Rom, Figure 88-1. Philadelphia. Lippincott.
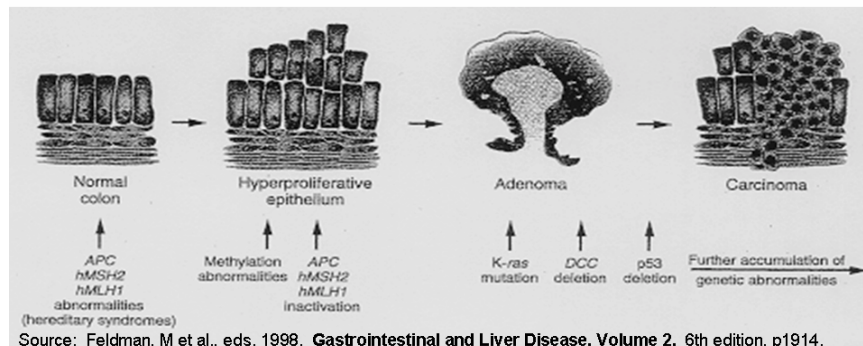
Source: Feldman, M et al., eds. 1998. **Gastrointestinal and Liver Disease, Volume 2,** 6th edition, p1914. Philadelphia, Saunders.

SLIDE 10 NOTES: STOOL DNA --Stool DNA is a provocative new technology. Its development was based on a number of observations about the natural history of colorectal cancer.

The progression of normal colonic mucosa through a series of mutational events that eventually leads to dysplastic and later neoplastic tissue is well-described. Dysplastic tissues exfoliate intact viable colonocytes. Those colonocytes have a representative genetic signature, and may manifest some of these abnormalities.

The observation of this process has led to studies attempting to isolate either individual or combined sets of tumor markers from stool in order to define the risk of colorectal cancer in populations. Most studies to date have focused on groups with known colorectal cancer or those at higher risk.

SLIDE 11

Test Performance
Stool DNA

Stool DNA	Polyps 6-9 mm	Polyps > 1 cm	Cancer
Sensitivity	10-82%		52-91%
Specificity	95%		90-100%

- About 50% of CRCs have K-*ras* mutations
  - False positives in IBD and lymphocytic colitis
- New multi-target assays have been developed
  - Uses: K-*ras*, APC, p53, BAT-26, L-DNA/DIA
  - Improved test performance over single mutations
    - Sensitivity in high-risk groups
      - Adenomas → 47-82%
      - Cancers → 64-91%
  - Screening populations (n=2,507)
    - Sensitivity for advanced adenomas → 15%
    - Specificity in patients with polyps → 95%

11

SLIDE 11 NOTES: It has been determined from patients with proven CRC that a number of mutations can be detected in cells obtained from stool specimens that correspond to mutations in the neoplasm in those affected individuals. Unfortunately, no individual mutation is found in all colorectal neoplasms. Consequently, researchers have assessed the utility of individual markers or multi-target assays. Some of the most common mutations in current assays include K-ras, APC, p53, BAT-26 and L-DNA (or DIA). An additional marker that has been studied is the D-Gal-B (1à3)D-GalNAc tumor marker (Sakamoto et al., 1993).

One recent study used a multi-target assay in order to assess test characteristics in patients with advanced neoplasia, and showed this assay had sensitivities of 64 percent and 57 percent for cancers and advanced adenomas, respectively (Tagore et al., 2003). Another study using similar multi-target assay also showed a sensitivity ranging from 73-82 percent for adenomas and 91 percent for cancers (Ahlquist et al., 2000). In a feasibility study, BAT-26 alone showed a sensitivity of 37 percent and a specificity of 100 percent (Traverso et al., 2002). K-ras has been assessed in a number of studies and its sensitivity as a sole marker ranges from 18-80 percent, but there are still some problems with non-neoplastic tissues manifesting stool positive for K-ras (specificity of 76 percent for organic disease in the colon) (Ahlquist et al., 2002; Villa et al., 1996).

Although sensitivity for adenomas of multi-target marketer tests has been on the order of 50 to 80 percent in high-risk groups, and even higher than that for cancers, a recent industry-sponsored study presented to the American College of Gastroenterology found sensitivity for advanced adenomas of only 15 percent with specificity of 95 percent.

To summarize, this is an exciting area that deserves further investigation, but is not yet ready for widespread screening of the population for CRC (Ahlquist, 2000).
Definitions: DIA = DNA Integrity Assay. L-DNA = "Long" DNA or High Molecular Weight DNA.

SLIDE 12

Test Performance
Virtual Colonoscopy

V.C.	Polyps 6-9 mm	Polyps > 1 cm	Cancer
Sensitivity	41-94%	32-96%	?-100%
Specificity	63-95%	74-98%	?-100%

- Evolving technology
  - Wide range of results
    - 2-D vs. 3-D
    - Software and hardware
    - Radiologist experience
    - Prevalence of population
  - Sensitivity for polyps
    - > 10 mm → 32-96%
    - 6-9 mm → 41-94%
  - Specificity for polyps
    - >10 mm → 74-98%
    - 6-9 mm → 63-95%

12

SLIDE 12 NOTES: Virtual colonoscopy is the latest technology to receive high visibility in the press. It is important to recognize, though, that this is still an evolving technology. As such, there has been a broad range of results (Gluecker et al., 2002; Johnson et al., 2003a; Laghi et al., 2002; Pescatore et al., 2000; Pickhardt et al., 2003; Pineau et al., 2003; Spinzi et al., 2001; Wong et al., 2002; Yee et al., 2001). That may be because of the range of technologies used, focusing on 2-D imaging, a combination of 2-D and 3-D, or 3-D imaging only, and a variety of those modalities at that.

The most recent publication provided the most impressive results to date, but these results have yet to be replicated (Pickhardt et al., 2003). The by-patient sensitivity for patients with polyps 10 mm and larger ranges from 75 percent to 100 percent and the by-polyp sensitivity for polyps 10 mm and larger ranges from 50 percent to 100 percent, with specificity ranging from 90-100 percent (Dachman, 2002). The inter-observer reliability appears to range from poor to very good, with kappa statistics ranging from -

0.56 to 0.89. ((Gluecker et al., 2002; Laghi et al., 2002; Pickhardt et al., 2003; Pescatore et al., 2000; Johnson et al., 2003b). Also important, the incidence of significant extra-colonic findings is 4.5-13 percent, which will impact the cost and cost-effectiveness of this test (Dachman, 2002; Kiesslich et al., 2001; Pickhardt et al., 2003).

In summary, for polyps, there is a broad range of sensitivity, and a narrower range for specificity. However, the sensitivity and specificity for cancer is approaching 100 percent if all of the studies are taken into account.

SLIDE 13

Virtual Colonoscopy Studies
All with n > 100

Author	Publication	Population	Technology	Contrast	Gold Standard	Kappa
Laghi, et al. (n=165)	*Am J Surgery* 2002	High-Risk~ Known Lesions	2-D Multiview 3-D Luminal	Typical OC prep CO_2	Optical Colonoscopy	Consensus Reading
Fenlon, et al. (n=100)	*NEJM* 1999	High-risk Screening	2-D Axial 3-D Fly-Through	Typical OC prep Air insufflation	Optical Colonoscopy	Consensus Reading
Yee, et al. (n=300)	*Radiology* 2001	Symptomatic Evaluations	2-D Multiview 3-D Fly-Through	Typical OC prep CO_2	Optical Colonoscopy	Consensus Reading
Pineau, et al. (n=205)	*Gastroenterology* 2003	Mixed Screening High-Risk	2-D Multiview 3-D Luminal	Typical OC prep CO_2	SU Optical Colonoscopy	Single Reader
Johnson, et al. (n=703)	*Gastroenterology* 2003	High-Risk Screening	2-D Assessment 3-D Confirmation	Gastroview CO_2	Optical Colonoscopy	-0.67-0.89 3 Readers
Pickhardt, et al. (n=1233)	*NEJM* 2003	Average-Risk Screening	3-D Fly-Through 2-D Confirmation	Barium DM/DS	SU Optical Colonoscopy	0.75-0.79 2 Readers

13

SLIDE 13 NOTES: Here is a list of the six studies with a sample size of 100 or greater. All have been reported in the recent past, before the conduct of this workshop.

Note that the populations studied are varied. Most of these studies have been in high-risk individuals with known lesions, or represent screening in high-risk populations. The only study that examined the use of virtual colonoscopy in average-risk individuals is the study reported recently in the New England Journal of Medicine by Pickhardt et al.

Several factors affect the results. The first is the use (or non-use) of contrast materials to enhance the image. Most of the studies reported on VC procedures using typical optical colonoscopy preparations of either air or CO2 insufflation. Two studies used added contrast to help subtract stool material from the image.

Second, the majority of studies used general optical colonoscopy as the gold standard for a true positive result. A more unbiased assessment of overall performance, however, is achieved by the use of segmental unblinding, which requires colonoscopists to re-examine patients for lesions found only on VC.

Third, measurement of inter-rater reliability is important in understanding the performance of VC. Most of the studies did not examine this aspect, because they had either a single reader or consensus readings.

SLIDE 14

V.C. Study Results
Per Patient Analysis

		Laghi et al. (n=165) *2	Fenlon et al. (n=100) *2	Yee et al. (n=300) *2	Pineau et al. (n=205) *1	Johnson et al. (n=703) *3	Pickhardt et al. (n=1233) *3
Sensitivity	<6	50%		82 %			
	6-9	82%	94 %	93 %	84.4 %	41-69% /65%	88.7 - 93.9 %
	>9	92%	96 %	100 %	90.0 %	35-72% /64%	93.8 %
	Overall	78%	82 %	93.9%	61.8 %		
Specificity	<6						
	6-9		92 %		83.1 %	88-95% /86%	79.6 - 94.9 %
	>9		96 %		94.6 %	97-98% /95%	96.0 %
	Overall	97%	84 %	56.6&72 %	70.7 %		
PPV	<6						
	6-9		92 %		58.5 %		
	>9		96 %	81%	64.3 %		
	Overall		82 %	62.9%	61.8 %		
NPV	<6						
	6-9		94 %		95.0 %		
	>9		96 %	97.2%	98.9 %		99%
	Overall		84 %	92.2%	70.7 %		

(*1) 1 reader   (*2) 2 readers   (*3) 3 readers   14

SLIDE 14 NOTES: Looking at the per-patient data gives the possibility of determining what will happen to patients who are subjected to screening by this modality.

Across all six studies, the sensitivity for small polyps has not been well studied. However, for intermediate polyps (between 6 and 9 mm), the sensitivity is probably 85 percent or more. (The Johnson study is an outlier for a variety of reasons.) Sensitivity for large polyps is generally 85 to 90 percent, but it is over 90 percent in most of the studies.

Specificity for intermediate-sized polyps has been 85-95 percent, and specificity for large polyps has been about 95 percent. The positive predictive value ranged from 60 to 90 percent, and the negative predictive value in general has been between 95 and 99 percent.

SLIDE 15

The Pickhardt Study -Methods

- Multi-center Enrollment
  - Adults ages 50-79
  - Adults with FH CRC ages 40-79
- Preparation
  - Fleet PhosphoSoda & Bisacodyl
  - 500 ml barium with clear diet
    - Solid-stool tagging
  - 120 mL of diatrozoate meglumine and diatrozoate sodium
    - Liquid opacification

15

SLIDE 15 NOTES: The Pickhardt Study -- In my opinion, The Pickhardt study (Pickhardt et al., 2003) is the one against which many others will be assessed in the future. This was a multi-center enrollment study. Average-risk individuals, mostly men, and a few high-risk individuals as defined by family history. (Of 1,233 patients, 1,201 were average-risk.)

The preparation was typical of optical colonoscopy, but barium was added for solid stool tagging, and diatrozoate meglumine and diatrozoate sodium were used to subtract liquids from the image. Diagnostic studies were performed in a typical fashion, but they used a rectal catheter with patient-controlled insufflation, looked at supine and prone images, and used high-speed GE technology. The reconstruction interval was small. Proprietary software was used to generate the 3D fly-though image, and 2D images were used only for confirmation.

Polyps were measured with electronic calipers. Extra-colonic findings were examined. And, all of the readers had read between 25 and 50 VC procedures via this technology before entering the study.

Optical colonoscopy was performed in a standard manner by 17 endoscopists, but as one of the endoscopists in the trial I am aware that the skill level varied among the 17. This study also had a stringent set of criteria for exclusion of individuals from randomization, most notably if they had a history of previous colonic studies or of any signs of cancer or polyps.

SLIDE 16

The Pickhardt Study –Diagnostic Methods

- Virtual Colonoscopy
  - Rectal catheter
    - Patient-controlled insufflation
    - Breath-holding
      - Supine and prone
  - GE Lightspeed or Ultra CT
    - Reconstruction interval 1 mm
    - Viatronix 3D Colon 1.2
      - Virtual 'fly-through'
      - Correlation with 2-D images
    - Polyps measured by electronic calipers
    - Extra-colonic findings recorded
  - Analyzed by 'experienced' radiologists

- Optical Colonoscopy
  - Standard video-endoscope
    - 17 'experienced' endoscopists
      - 14 GI, 3 GS
    - Polyps measured by calibrated linear probe
    - Segmental unblinding
    - Polyps resected
      - Sent for pathology
    - Procedural times recorded
    - Questionnaires provided

16

SLIDE 16 NOTES: No notes.

SLIDE 17

The Pickhardt Study -Statistical Analysis

- OC = 'Reference Standard'
- Primary outcome
  - Polyps > 6 mm
- "Advanced neoplasia"
  - Polyps > 10 mm
  - HGD, TV, or cancer
- Polyp Matching VC/OC
  - True positive
    - Same/adjacent segment
    - Diameters equal +/- 50%
  - True positive by size
    - Polyp of $\geq$ size on VC & OC

- Analysis using:
  - McNemar's test
  - Fisher's exact test
  - Chi-square tests
  - Paired t-tests
- Interobserver reliability
  - 100 random VCs
    - 2nd read by blinded radiologist
    - Kappa defined

17

SLIDE 17 NOTES: The Pickhardt study employed segmental unblinding, which means that as the colonoscope passes through a segment, any indication from the VC results that a lesion may have been missed would lead to re-examination of that segment of the colon to attempt to verify its existence.

Polyps were matched between VC and OC by a finding of polyp in the same or adjacent segment with diameters that were within 50 percent of one another.

All polyps found on optical colonoscopy were resected and sent for pathology.

Procedure times were recorded, and questionnaires were provided to patients. The reference standard for a true positive was the segmentally unblinded optical colonoscopy result.

SLIDE 18

Pickhardt Results

- 1233/1253 patients completed VC/OC
  - 1201 average risk
- Adenomas
  - > 10 mm→ 3.9%
  - > 8 mm → 6.7%
  - > 6 mm → 13.6%
  - Malignant → 0.4%
  - Advanced polyps ≤ 5 mm → 0.1%

Table 2. Distribution of Adenomatous and Nonadenomatous Polyps According to Size and Location in 1233 Asymptomatic Adults.*

Variable	Polyp Size			
	≤5 mm	6–9 mm	≥10 mm	Any
	number of polyps			
Rectum				
Adenomatous	28	17	10	55
Nonadenomatous	167	24	4	195
Sigmoid colon				
Adenomatous	75	46	7	128
Nonadenomatous	192	31	8	231
Descending colon				
Adenomatous	39	18	6	63
Nonadenomatous	51	8	1	60
Splenic flexure				
Adenomatous	11	5	2	18
Nonadenomatous	27	5	2	34
Transverse colon				
Adenomatous	47	29	3	79
Nonadenomatous	60	5	4	69
Hepatic flexure				
Adenomatous	29	8	0	37
Nonadenomatous	28	5	1	34
Ascending colon				
Adenomatous	73	23	16	112
Nonadenomatous	60	20	8	88
Cecum				
Adenomatous	42	13	7	62
Nonadenomatous	37	5	3	45
Total				
Adenomatous	344	159	51	554
Nonadenomatous	622	103	31	756

* A total of 611 of the 1233 patients had no polyps. Two polyps, located in the cecum and the ascending colon, were malignant.

SLIDE 18 NOTES: No notes.

SLIDE 19

Pickhardt Results – Test Performance

- Polyps on VC not OC
  - 55 polyps
    - 21 TAs > 6 mm
- "Advanced Neoplasms"
  - Sensitivity
    - VC → 91.5%
    - OC → 88.1%
- Colon Cancer
  - Sensitivity
    - VC → 100% (2/2)
    - OC → 50% (1/2)

SLIDE 19 NOTES: To summarize comparative test performance, VC found 55 polyps not seen on OC, and 21 of those were tubular adenomas greater than 6 mm. The sensitivity for advanced neoplasia was 91.5 percent for VC and 88 percent for colonoscopy. The difference was not statistically significant.

SLIDE 20

Pickhardt Results

- **Extra-colonic findings (rates are half of previous reports)**
  - Potentially high clinical importance → 4.5%
    - 7/56 had extra-colonic cancer or AAA (12.5%)
  - Potentially moderate clinical importance → 13.5%
    - Nephrolithiasis (7.9%) or cholelithiasis (5.6%)
- **Procedural time**
  - VC → 14.1 minutes (in CT suite)
  - OC → 31.5 minutes (64.4 additional min.s in recovery)
- **Satisfaction (81.5% questionnaires returned)**
  - Greater discomfort: VC 54.3% vs. OC 38.1%
  - More acceptable(convenience): VC 68.3% vs. OC 24.1%

20

SLIDE 20 NOTES: The additional outcome measures described in this slide – extra-colonic findings, procedural time, and patient satisfaction – are secondary outcomes that may be important in evaluating the cost-effectiveness of this technology.

SLIDE 21

Pickhardt Results
Summary Statistics

- VC for polyps > 8 mm
  - Sensitivity & Specificity > 90%
  - 1 in 6 patients will get OC
  - Kappa > 0.75
  - VC = OC for CT/radiologist time
    - Quicker for patients
  - VC less comfortable than OC
    - But more acceptable
- Bottom-line:
  - VC better than OC for detection of adenomas

21

SLIDE 21 NOTES: Overall summary statistics for VC performance --with a cutoff range of 8 mm or greater, sensitivity and specificity of VC are greater than 90 percent. However, test positivity rates are high, which implies that 1 in 6 patients would be referred for optical colonoscopy.

The bottom line from this study is that VC is equivalent or better than OC for finding adenomas.

SLIDE 22

Pickhardt Study: Questions/Issues

- What is the gold standard in this study?
  - "OC was the reference standard"
    - But sensitivity defined for VC <u>and</u> OC. Perhaps this is reasonable, but...
- Proceduralists → Staff radiologists vs. Staff/Fellows GI/Surgeons
- What about data for hyperplastic polyps?
  - Specificity of VC for any polyps
    - Polyps > 10 mm → 97.4%   ⎫ VC HPs 5-8 mm would get VC q 2-3 years
    - Polyps > 8 mm → 95.0%   ⎬ All OC HPs would get next OC at 10 years
    - Polyps > 6 mm → 84.5%   ⎭
- What about small advanced polyps?
  - Only 0.1% polyps ≤ 5 mm advanced
    - But what about with 6 or 7 mm polyps?
  - Test positive rate at 8 mm is 13.5% (1 in 6 get colonoscopy)
    - At 6 mm, 1 in every 3 patients get colonoscopy (where spec = 80%)

8

SLIDE 22 NOTES: No notes.

SLIDE 23

Test Performance
Other Studies

- FDG-PET
- Dark lumen MR colography
- Chromoendoscopy

23

SLIDE 23 NOTES: Here is a brief description of some provocative technologies 'not (yet) ready for prime time:'

Flurodioxyglucose Positron Emission Tomography (FDG-PET) FDGs take up metabolically active cells. In a recently published study (Drenth, et al., 2001)., this technology achieved a sensitivity of 74 percent and specificity of 84 percent for polyps greater than 3 mm or cancers throughout the colonic tract.

Dark lumen MR colography (Lubolt et al., 1998; Morrin et al., 2001) probably has characteristics similar to VC based on CT, but that technology has been less well studied. One preliminary study showed very high sensitivity for large polyps and lower sensitivity for polyps in the intermediate range (Lubolt et al., 1998).

Chromoendoscopy (Kiesslich et al., 2001; Lee et al., 2003) – the application of indigocarmine dye through a colonoscope --may actually improve the ability to analyze tissue characteristics during the endoscopic procedure. However, that capability may come at the expense of increased procedure cost and time.

SLIDE 24

Test Performance
The Bottom-Line

Test	Sensitivity Polyps	Specificity Polyps	Sensitivity Cancer	Specificity Cancer
FOBT	5-80%	20-97%	20-40%	68-100%
Rehydrated FOBT	5-90%	35-100%	40-65%	72-100%
Immunochemical FOBT	4-89%	62-98%	20-100%	63-100%
Barium Enema	48-100%	67-100%	62-100%	90-100%
Flexible Sigmoidoscopy	33-78%	84%	67-90%	85-95%
Colonoscopy	75-100%	90-98%	79-100%	90-96%
Virtual Colonoscopy	32-96%	63-98%	90-100%	90-100%
Stool DNA	10-82%	95%	52-91%	90-100%

24

SLIDE 24 NOTES: This table shows the broad range of sensitivity and specificity for polyps and cancer across all of the technologies reviewed here.

SLIDE 25

Test Performance
The Bottom-Line

Test	Sensitivity Polyps	Specificity Polyps	Sensitivity Cancer	Specificity Cancer
FOBT	< 30%	50%	~40%	80⁺%
Rehydrated FOBT				
Immunochemical FOBT				
Barium Enema				
Flexible Sigmoidoscopy				
Colonoscopy				
Virtual Colonoscopy	90⁺%	95%	~100%	~100%
Stool DNA	10-82%	95%	52-91%	90-100%

25

SLIDE 25 NOTES: No notes.

SLIDE 26

Summary Comments

- No test has been established as the 'Gold Standard'
- Tests
  - FOBTs and IFOBTs
    - Wide range of sensitivity and specificity
      - Lower sensitivity, but has demonstrated benefit
  - Barium Enema
    - Has acceptable sensitivity and specificity
      - Suffers from limited usage, inadequate study, inconvenience
  - Sigmoidoscopy
    - Has acceptable sensitivity and specificity
      - Limited exam of colon, usage/training suffers colonoscopy's role

26

SLIDE 26 NOTES: To summarize the entire literature, I would say that at present we have no clear gold standard for overall analysis. However, the segmentally unblinded colonoscopy using complementary studies is probably the best approach to defining the gold standard.

FOBTs and IFOBTs have a wide range of sensitivity, but even at low sensitivity, they have demonstrated a benefit in reducing colorectal cancer.

Although barium enema has acceptable sensitivity and specificity, it is in my opinion a technology on the wane, with fewer people trained to perform the procedure.

Colonoscopy has high sensitivity and specificity, but it is impacted by a significant miss rate, incomplete examinations and complications.

Virtual colonoscopy still has a broad range of sensitivity and specificity, and practical issues such as the hardware and software used, the inter-observer reliability, and radiologist training has to be addressed before it is ready for prime time.

SLIDE 27

Summary Comments

- Tests
    - Colonoscopy
        - Has high sensitivity and specificity (?)
            - Miss rate, incomplete examinations and complications
    - Virtual Colonoscopy
        - Has a range of sensitivity and specificity (high specificity)
            - Hard-/Software, inter-observer reliability, inability for intervention
    - Stool DNA
        - Has a range of sensitivity, specificity
            - Limited number of studies, assay dependent, lab dependent

27

SLIDE 27 NOTES: No notes.

REFERENCES

Adamsen, S. and O. Kronborg, Acceptability and compliance in screening for colorectal cancer with fecal occult blood test. *Scand J Gastroenterol*, 1984. 19(4): p. 531-4.

Ahlquist, D.A. and A.P. Shuber, Stool screening for colorectal cancer: evolution from occult blood to molecular markers. *Clinica Chimica Acta*, 2002. 315(1-2): p. 157-68.

Ahlquist, D.A., J.E. Skoletsky, K.A. Boynton, J.J. Harrington, D.W. Mahoney, W.E. Pierceall, S.N. Thibodeau, and A.P. Shuber, Colorectal cancer screening by detection of altered human DNA in stool: feasibility of a multitarget assay panel. *Gastroenterol*, 2000. 119(5): p. 1219-27.

Ahlquist, D.A., Molecular stool screening for colorectal cancer. Using DNA markers may be beneficial, but large scale evaluation is needed. *BMJ*, 2000. 321(7256): p. 254-5.

Beeker, C., J.M. Kraft, B.G. Southwell, et al., Colorectal cancer screening in older men and women: qualitative research findings and implications for intervention. *J Community Health*, 2000. 25(3): p. 263-78.

Bond, J.H., Colorectal Cancer Update. Prevention, Screening, Treatment and Surveillance for High-Risk Groups. *Med Clin N Am*, 2000. 84: p. 1163-82.

Bouvier, V., C. Herbert, H. Lefevre, and G. Launoy, Stage of extension and treatment for colorectal cancer after a negative test and among non-responders in mass screening with guaiac faecal occult blood test: a French experience. *Eur J Cancer Prev*, 2001. 10(4): p. 323-326.

Boyle, P., D.G. Zaridze, and M. Smans, Descriptive epidemiology of colorectal cancer. *Int J Cancer*, 1985. 36: p. 9-15.

Brenes, G.A. and E.D. Paskett, Predictors of stage of adoption for colorectal cancer screening. *Prev Med*, 2000. 31(4): p. 410-6.

Brenner, H., V. Arndt, T. Sturmer, C. Stegmaier, H. Ziegler, and G. Dhom, Long-lasting reduction of risk of colorectal cancer following screening endoscopy. *Br J Cancer*, 2001. 85(7): p. 972-6.

Brevinge, H., E. Lindholm, S. Buntzen, and J. Kewenter, Screening for colorectal neoplasia with faecal occult blood testing compared with flexible sigmoidoscopy directly in a 55-56 years' old population. *Int J Colorectal Dis*, 1997. 12(5): p. 291-5.

Brewster, N.T., D.C. Grieve, and J.H. Saunders, Double-contrast barium enema and flexible sigmoidoscopy for routine colonic investigation. *Br J Surg*, 1994. 81(3): p. 445-7.

Byrne, M.F., Primary Screening With Colonoscopy for Colorectal Cancer: A Targeted Algorithm? *Am J Gastroenterol*, 2003. 98(12): p. 2587-8.

Chu, K.C., R.E. Tarone, W.H. Chow, B.F. Hankey, and L.A. Ries, Temporal patterns in colorectal cancer incidence, survival, and mortality from 1950 through 1990. [see comment]. *J Natl Cancer Inst*, 1994. 86(13): p. 997-1006.

Cirocco, W.C. and L.C. Rusin, Factors that predict incomplete colonoscopy. *Dis Colon Rect*, 1995. 38(9): p. 964-8.

Cooper, G.S., Z. Yuan, A. Chak, and A.A. Rimm, Patterns of endoscopic follow-up after surgery for nonmetastatic colorectal cancer. *Gastrointest Endosc*, 2000. 52(1): p. 33-8.

Dachman, A.H., Diagnostic performance of virtual colonoscopy. *Abdom Imaging*, 2002. 27(3): p. 260-7.

de Zwart, I.M., G. Griffioen, M.P. Shaw, C.B. Lamers, and A. de Roos, Barium enema and endoscopy for the detection of colorectal neoplasia: sensitivity, specificity, complications and its determinants. *Clin Radiol*, 2001. 56(5): p. 401-9.

Drenth, J.P., F.M. Nagengast, and W.J. Oyen, Evaluation of (pre-)malignant colonic abnormalities: endoscopic validation of FDG-PET findings. *Eur J Nucl Med*, 2001. 28(12): p. 1766-9.

Ederer, F., T.R. Church, and J.S. Mandel, Fecal occult blood screening in the Minnesota study: role of chance detection of lesions.[see comment]. *J Natl Cancer Inst*, 1997. 89(19): p. 1423-8.

Eisen, G.M., C.Y. Kim, D.E. Fleischer, R.A. Kozarek, D.L. Carr-Locke, T.C. Li, C.J. Gostout, S.J. Heller, E.A. Montgomery, F.H. Al-Kawas, J.H. Lewis, and S.B. Benjamin, High-resolution chromoendoscopy for classifying colonic polyps: a multicenter study. *Gastrointest Endosc*, 2002. 55(6): p. 687-94.

Erban, S., J. Zapka, E. Puleo, and M. Vickers-Lahti, Colorectal cancer screening in Massachusetts: measuring compliance with current guidelines. *Eff Clin Pract*, 2001. 4(1): p. 10-7.

Favennec, L., N. Kapel, D. Meillet, C. Chochillon, and J.G. Gobert, Detection of occult blood in stools: comparison of three gaiac tests and a latex agglutination test. [erratum appears in *Ann Biol Clin (Paris)* 1992. 50(8):618]. *Ann Biol Clin (Paris)*, 1992. 50(5): p. 311-3.

Gann, P.H., J.E. Manson, R.J. Glynn, et al., Low-dose aspirin and incidence of colorectal tumors in a randomized trial. *J Natl Cancer Inst*, 1993. 85(15): p. 1220-24.

Giovanucci, E., E.B. Rimm, M.J. Stampfer, et al., Aspirin use and the risk of colorectal cancer and adenoma in male health professionals. *Ann Intern Med*, 1994. 121: p. 241-8.

Glick, S., J.L. Wagner, and C.D. Johnson, Cost-effectiveness of double-contrast barium enema in screening for colorectal cancer. *AJR. Am J Roengenol*, 1998. 170(3): p. 629-36.

Gluecker, T., G. Dorta, W. Keller, P. Jornod, R. Meuli, and P. Schnyder, Performance of multidetector computed tomography colonography compared with conventional colonoscopy. *Gut*, 2002. 51(2): p. 207-11.

Gopalswamy, N., H.P. Stelling, R.J. Markert, H.N. Maimon, S.D. Wahlen, and R.I. Haddy, A comparative study of eight fecal occult blood tests and HemoQuant in patients in whom colonoscopy is indicated. *Arch Fam Med*, 1994. 3(12): p. 1043-8.

Greenberg, P.D., L. Bertario, R. Gnauck, O. Kronborg, J.D. Hardcastle, M.S. Epstein, D. Sadowski, R. Sudduth, G.R. Zuckerman, and D.C. Rockey, A prospective multicenter evaluation of new fecal occult blood tests in patients undergoing colonoscopy. *Am J Gastroenterol*, 2000. 95(5): p. 1331-8.

Greenlee, R.T., M.B. Hill-Harmon, T. Murray, et al., Cancer Statistics, 2001. *CA Cancer J Clin*, 2001(51): p. 15-36.

Hixson, L.J., M.B. Fennerty, R.E. Sampliner, and H.S. Garewal, Prospective blinded trial of the colonoscopic miss-rate of large colorectal polyps. *Gastrointest Endosc*, 1991. 37(2): p. 125-7.

Hixson, L.J., M.B. Fennerty, R.E. Sampliner, D. McGee, and H. Garewal, Prospective study of the frequency and size distribution of polyps missed by colonoscopy. *J Natl Cancer Inst*, 1990. 82(22): p. 1769-72.

Hope, R.L., G. Chu, A.H. Hope, R.G. Newcombe, P.E. Gillespie, and S.J. Williams, Comparison of three faecal occult blood tests in the detection of colorectal neoplasia. *Gut*, 1996. 39(5): p. 722-5.

Irvine, E.J., J. O'Connor, R.A. Frost, P. Shorvon, S. Somers, G.W. Stevenson, and R.H. Hunt, Prospective comparison of double contrast barium enema plus flexible sigmoidoscopy v colonoscopy in rectal bleeding: barium enema v colonoscopy in rectal bleeding. [comment]. *Gut*, 1988. 29(9): p. 1188-93.

Itoh, M., K. Takahashi, H. Nishida, K. Sakagami, and T. Okubo, Estimation of the optimal cut off point in a new immunological faecal occult blood test in a corporate colorectal cancer screening programme. *J Med Screen*, 1996. 3(2): p. 66-71.

Jeanson, A., J. Jamart, J.M. Maisin, R. Vanheuverzwyn, P. Gohy, J.C. Debongnie, and B. Rentier, Assessment of the new immunological test Hemoblot for detecting occult blood in faeces. *Eur J Cancer Prev*, 1994. 3(5): p. 407-12.

Jensen, J., J. Kewenter, M. Asztely, G. Lycke, and J. Wojciechowski, Double contrast barium enema and flexible rectosigmoidoscopy: a reliable diagnostic combination for detection of colorectal neoplasm. *Br J Surg*, 1990. 77(3): p. 270-2.

Johnson, C.D., A.Y. Toledano, B.A. Herman, A.H. Dachman, E.G. McFarland, M.A. Barish, J.A. Brink, R.D. Ernst, J.G. Fletcher, R.A. Halvorsen, Jr., A.K. Hara, K.D. Hopper, R.E. Koehler, D.S. Lu, M. Macari, R.L. Maccarty, F.H. Miller, M. Morrin, E.K. Paulson, J. Yee, M. Zalis, and American College of Radiology Imaging Network, Computerized tomographic colonography: performance evaluation in a retrospective multicenter setting. *Gastroenterol*, 2003. 125(3): p. 688-95.

Johnson, C.D., H.C. Carlson, W.F. Taylor, and L.P. Weiland, Barium enemas of carcinoma of the colon: sensitivity of double- and single-contrast studies. *AJR Am J Roentgenol*, 1983. 140(6): p. 1143-9.

Johnson, C.D., W.S. Harmsen, L.A. Wilson, R.L. Maccarty, T.J. Welch, D.M. Ilstrup, and D.A. Ahlquist, Prospective blinded evaluation of computed tomographic colonography for screen detection of colorectal polyps. [comment]. *Gastroenterol*, 2003. 125(2): p. 311-9.

Kewenter, J., H. Brevinge, B. Engaras, and E. Haglind, The yield of flexible sigmoidoscopy and double-contrast barium enema in the diagnosis of neoplasms in the large bowel in patients with a positive Hemoccult test. [see comment]. *Endoscopy*, 1995. 27(2): p. 159-63.

Kewenter, J., S. Bjork, E. Haglind, L. Smith, J. Svanvik, and C. Ahren, Screening and rescreening for colorectal cancer. A controlled trial of fecal occult blood testing in 27,700 subjects. *Cancer*, 1988. 62(3): p. 645-51.

Khullar, S.K. and J.A. DiSario, Colon cancer screening. Sigmoidoscopy or colonoscopy. *Gastrointest Endosc Clin N Am*, 1997. 7(3): p. 365-86.

Kiesslich, R., et al., Chromoendoscopy with indigocarmine improves the detection of adenomatous and nonadenomatous lesions in the colon. [see comment]. *Endoscopy*, 2001. 33(12): p. 1001-6.

Ko, C.W., J.A. Dominitz, and T.D. Nguyen, Fecal occult blood testing in a general medical clinic: comparison between guaiac-based and immunochemical-based tests. *Am J Med*, 2003. 115(2): p. 111-4.

Kronborg, O., Colonic screening and surveillance. *Best Practice & Research in Clinical Gastroenterology*, 2001. 15(2): p. 301-16.

Laghi, A., R. Iannaccone, I. Carbone, C. Catalano, V. Panebianco, E. Di Giulio, A. Schillaci, and R. Passariello, Computed tomographic colonography (virtual colonoscopy): blinded prospective comparison with conventional colonoscopy for the detection of colorectal neoplasia. *Endoscopy*, 2002. 34(6): p. 441-6.

Landis, S.H., T. Murray, S. Bolden, et al., Cancer Statistics. *CA Cancer J Clin*, 1999. 49: p. 8-31.

Lee, J.H., et al., Detection of colorectal adenomas by routine chromoendoscopy with indigocarmine. *Am J Gastroenterol*, 2003. 98(6): p. 1284-8.

Lee, J.H., J.W. Kim, Y.K. Cho, C.I. Sohn, W.K. Jeon, B.I. Kim, and E.Y. Cho, Detection of colorectal adenomas by routine chromoendoscopy with indigocarmine. *Am J Gastroenterol*, 2003. 98(6): p. 1284-8.

Leinicke, J.L., W.J. Dodds, W.J. Hogan, and E.T. Stewart, A comparison of colonoscopy and roentgenography for detecting polypoid lesions of the colon. *Gastrointest Radiol*, 1977. 2(2): p. 125-8.

Lichtenstein, G.R., P.D. Park, W.B. Long, G.G. Ginsberg, and M.L. Kochman, Use of a push enteroscope improves ability to perform total colonoscopy in previously unsuccessful attempts at colonoscopy in adult patients. *Am J Gastroenterol*, 1999. 94(1): p. 187-90.

Lieberman, D., Cost-effectiveness of screening colorectal cancer in the general population. *Gastrointest Endosc*, 2001. 54(4): p. 537-8.

Lieberman, D.A., D.G. Weiss; Veterans Affairs Cooperative Study, One-time screening for colorectal cancer with combined fecal occult-blood testing and examination of the distal colon. [see comment][summary for patients in *Gastrointest Endosc*, 2002 May; 55(6): p. 760-1]. *N Engl J Med*, 2001. 345(8): p. 555-60.

Loose, H.W. and C.B. Williams, Barium enema versus colonoscopy. *Proc R Soc Med*, 1974. 67(10): p. 1033-6.

Luboldt, W., P. Steiner, P. Bauerfeind, P. Peklonen, and J.F. Debatin, Detection of mass lesions with MR colonography: preliminary report. *Radiology*, 1998. 207: p. 59-65.

Mandel, J.S., J.H. Bond, T.R. Church, et al., Reducing mortality from colorectal cancer by screening for fecal occult blood. Minnesota Colon Cancer Control Study. *N Engl J Med*, 1993. 328(19): p. 1365-74.

Markus, J.B., S. Somers, B.P. O'Malley, and G.W. Stevenson, Double-contrast barium enema studies: effect of multiple reading on perception error. *Radiology*, 1990. 175(1): p. 155-6.

Miyoshi, H., M. Oka, K. Sugi, O. Saitoh, K. Katsu, and K. Uchida, Accuracy of detection of colorectal neoplasia using an immunochemical occult blood test in symptomatic referred patients: comparison of retrospective and prospective studies. *Intern Med*, 2000. 39(9): p. 701-6.

Morrin, M.M., et al., MR colonography using colonic distention with air as the contrast material: work in progress. *AJR Am J Roentgenol*, 2001. 176(1): p. 144-6.

Nakama, H., B. Zhang, and X. Zhang, Evaluation of the optimum cut-off point in immunochemical occult blood testing in screening for colorectal cancer. *Eur J Cancer*, 2001. 37(3): p. 398-401.

Nakama, H., M. Yamamoto, N. Kamijo, T. Li, N. Wei, A.S. Fattah, and B. Zhang, Colonoscopic evaluation of immunochemical fecal occult blood test for detection of colorectal neoplasia. *Hepatogastroenterology*, 1999. 46(25): p. 228-31.

Nakama, H., N. Kamijo, K. Fujimori, A.S. Fattah, and B. Zhang, Relationship between fecal sampling times and sensitivity and specificity of immunochemical fecal occult blood tests for colorectal cancer: a comparative study. *Dis Colon Rectum*, 1997. 40(7): p. 781-4.

National Cancer Institute. http://www.cancer.org.

Nelson, D.B., K.R. McQuaid, J.H. Bond, D.A. Lieberman, D.G. Weiss, and T.K. Johnston, Procedural success and complications of large-scale screening colonoscopy. *Gastrointest Endosc*, 2002. 55(3): p. 307-14.

Newcomb, P.A., B.E. Storer, L.M. Morimoto, A. Templeton, and J.D. Potter, Long-term efficacy of sigmoidoscopy in the reduction of colorectal cancer incidence. [see comment]. *J Natl Cancer Inst*, 2003. 95(8): p. 622-5.

Newcomb, P.A., R.G. Norfleet, B.E. Storer, T.S. Surawicz, and P.M. Marcus, Screening sigmoidoscopy and colorectal cancer mortality. [see comment]. *J Natl Cancer Inst*, 1992. 84(20): p. 1572-5.

Niv, Y. and A.D. Sperber, Sensitivity, specificity, and predictive value of fecal occult blood testing (Hemoccult II) for colorectal neoplasia in symptomatic patients: a prospective study with total colonoscopy. *Am J Gastroenterol*, 1995. 90(11): p. 1974-7.

Obrecht, W.F., Jr., W.C. Wu, D.W. Gelfand, and D.J. Ott, The extent of successful colonoscopy: a second assessment using modern equipment. *Gastrointest Radiol*, 1984. 9(2): p. 161-2.

Osgard, E., J.L. Jackson, and J. Strong, A randomized trial comparing three methods of bowel preparation for flexible sigmoidoscopy. *Am J Gastroenterol*, 1998. 93(7): p. 1126-30.

Ott, D.J., W.C. Wu, and D.W. Gelfand, Extent of colonic visualization with the fiberoptic sigmoidoscope. *J Clin Gastroenterol*, 1982. 4(4): p. 337-41.

Painter, J., D.B. Saunders, G.D. Bell, C.B. Williams, R. Pitt, and J. Bladen, Depth of insertion at flexible sigmoidoscopy: implications for colorectal cancer screening and instrument design. [see comment]. *Endoscopy*, 1999. 31(3): p. 227-31.

Parikh, A., R. Ramamoorthy, K.H. Kim, B.K. Holland, and J. Houghton, Fecal occult blood testing in a noncompliant inner city minority population: increased compliance and adherence to screening procedures without loss of test sensitivity using stool obtained at the time of in- office rectal examination. *Am J Gastroenterol*, 2001. 96(6): p. 1908-13.

Parker, S.L., T. Tong, S. Bolden, et al., Cancer statistics, 1996. *CA Cance J Clin*, 1996. 46(1): p. 5-27.

Pescatore, P., T. Glucker, J. Delarive, R. Meuli, D. Pantoflickova, B. Duvoisin, P. Schnyder, A.L. Blum, and G. Dorta, Diagnostic accuracy and interobserver agreement of CT colonography (virtual colonoscopy). *Gut*, 2000. 47(1): p. 126-30.

Pickhardt, P.J., J.R. Choi, I. Hwang, J.A. Butler, M.L. Puckett, H.A. Hildebrandt, R.K. Wong, P.A. Nugent, P.A. Mysliwiec, and W.R. Schindler, Computed tomographic virtual colonoscopy to screen for colorectal neoplasia in asymptomatic adults. [see comment]. *N Engl J Med*, 2003. 349(23): p. 2191-2200.

Pignone, M., S. Saha, T. Hoerger, and J. Mandelblatt, Cost-effectiveness analyses of colorectal cancer screening: a systematic review for the U.S. Preventive Services Task Force. [comment][summary for patients in *Ann Intern Med*, 2002. 137(2): p. I38]. *Ann Intern Med*, 2002. 137(2): p. 96-104.

Pineau, B.C., E.D. Paskett, G.J. Chen, M.A. Espeland, K. Phillips, J.P. Han, C. Mikulaninec, and D.J. Vining, Virtual colonoscopy using oral contrast compared with colonoscopy for the detection of patients with colorectal polyps. [comment]. *Gastroenterol*, 2003. 125(2): p. 304-10.

Postic, G., D. Lewin, C. Bickerstaff, and M.B. Wallace, Colonoscopic miss rates determined by direct comparison of colonoscopy with colon resection specimens. *Am J Gastroenterol*, 2002. 97(12): p. 3182-5.

Provenzale, D., Cost-effectiveness of screening the average-risk population for colorectal cancer. *Gastrointest Endosc Clin N Am*, 2002. 12(1): p. 93-109.

Rex, D.K., C.S. Cutler, G.T. Lemmel, E.Y. Rahmani, D.W. Clark, D.J. Helper, G.A. Lehman, and D.G. Mark, Colonoscopic miss rates of adenomas determined by back-to-back colonoscopies. [see comment]. *Gastroenterol*, 1997. 112(1): p. 24-8.

Rex, D.K., Colonoscopic withdrawal technique is associated with adenoma miss rates. *Gastrointest Endosc*, 2000. 51(1): p. 33-6.

Rex, D.K., Colonoscopy: a review of its yield for cancers and adenomas by indication. *Am J Gastroenterol*, 1995. 90(3): p. 353-65.

Rex, D.K., E.Y. Rahmani, J.H. Haseman, G.T. Lemmel, S. Kaster, and J.S. Buckley, Relative sensitivity of colonoscopy and barium enema for detection of colorectal cancer in clinical practice. [comment]. *Gastroenterol*, 1997. 112(1): p. 17-23.

Rex, D.K., G.A. Lehman, J.C. Lappas, and R.E. Miller, Sensitivity of double-contrast barium study for left-colon polyps. *Radiology*, 1986. 158(1): p. 69-72.

Rex, D.K., Rationale for colonoscopy screening and estimated effectiveness in clinical practice. *Gastrointest Endosc Clin N Am*, 2002. 12(1): p. 65-75.

Rockey, D.C., J. Koch, J.P. Cello, L.L. Sanders, and K. McQuaid, Relative frequency of upper gastrointestinal and colonic lesions in patients with positive fecal occult-blood tests. *N Engl J Med*, 1998. 339(3): p. 153-9.

Rozen, P., J. Knaani, and N. Papo, Evaluation and comparison of an immunochemical and a guaiac faecal occult blood screening test for colorectal neoplasia. *Eur J Cancer Prev*, 1995. 4(6): p. 475-81.

Rozen, P., J. Knaani, and Z. Samuel, Comparative screening with a sensitive guaiac and specific immunochemical occult blood test in an endoscopic study. *Cancer*, 2000. 89(1): p. 46-52.

Rozen, P., J. Knaani, and Z. Samuel, Performance characteristics and comparison of two immunochemical and two guaiac fecal occult blood screening tests for colorectal neoplasia. *Dig Dis Sci*, 1997. 42(10): p. 2064-71.

Rozen, P., Prospects for the worldwide control of colorectal cancer through screening. *Gastrointest Endosc*, 2002. 55(6): p. 755-9.

Saito, Y., P. Slezak, and C. Rubio, The diagnostic value of combining flexible sigmoidoscopy and double-contrast barium enema as a one-stage procedure. *Gastrointest Radiol*, 1989. 14(4): p. 357-9.

Sakamoto, K., M. Muratani, T. Ogawa, and Y. Nagamachi, Evaluation of a new test for colorectal neoplasms: a prospective study of asymptomatic population. *Cancer Biother*, 1993. 8(1): p. 49-55.

Shehadeh, I., S. Rebala, R. Kumar, R.J. Markert, C. Barde, and N. Gopalswamy, Retrospective analysis of missed advanced adenomas on surveillance colonoscopy. *Am J Gastroenterol*, 2002. 97(5): p. 1143-7.

Shumaker, D.A., A. Zaman, and R.M. Katon, Use of a variable-stiffness colonoscope allows completion of colonoscopy after failure with the standard adult colonoscope. *Endoscopy*, 2002. 34(9): p. 711-4.

Sieg, A., A. Hertel, M.R. John, K. Luthgens, and H. Schmidt-Gayk, Screening for colorectal neoplasms with a new immunological human faecal haemoglobin and albumin test. *Eur J Cancer Prev*, 1998. 7(4): p. 279-85.

Sieg, A., C. Thoms, K. Luthgens, M.R. John, and H. Schmidt-Gayk, Detection of colorectal neoplasms by the highly sensitive hemoglobin-haptoglobin complex in feces. *Int J Colorectal Dis*, 1999. 14(6): p. 267-71.

Silverman, M.A., Cancer screening in the elderly population. *Hematol Oncol Clin North Am*, 2000. 14(1): p. 89-112.

Slattery, M.L., S.L. Edwards, K.M. Boucher, et al., Lifestyle and colon cancer: An assessment of factors associated with risk. *Am J Epidemiol*, 1999. 150: p. 869-77.

Sonnenberg, A., Cost-effectiveness in the prevention of colorectal cancer. *Gastroenterol Clin North Am*, 2002. 31(4): p. 1069-91.

Spinzi, G., G. Belloni, A. Martegani, A. Sangiovanni, C. Del Favero, and G. Minoli, Computed tomographic colonography and conventional colonoscopy for colon diseases: a prospective, blinded study. *Am J Gastroenterol*, 2001. 96(2): p. 394-400.

Steine, S., A. Stordahl, O.C. Lunde, K. Loken, and E. Laerum, Double-contrast barium enema versus colonoscopy in the diagnosis of neoplastic disorders: aspects of decision-making in general practice. *Fam Pract*, 1993. 10(3): p. 288-91.

Stewart, B.T., J.O. Keck, A.V. Duncan, N.M. Santamaria, and P. Allen, Difficult or incomplete flexible sigmoidoscopy: implications for a screening programme. [see comment]. *Aust N Z J Surg*, 1999. 69(1): p. 19-21.

Sung, J.J., F.K. Chan, W.K. Leung, J.C. Wu, J.Y. Lau, J. Ching, K.F. To, Y.T. Lee, Y.W. Luk, N.N. Kung, S.P. Kwok, M.K. Li, and S.C. Chung, Screening for colorectal cancer in Chinese: comparison of fecal occult blood test, flexible sigmoidoscopy, and colonoscopy. *Gastroenterol*, 2003. 124(3): p. 608-14.

Tagore, K.S., M.J. Lawson, J.A. Yucaitis, R. Gage, T. Orr, A.P. Shuber, and M.E. Ross, Sensitivity and specificity of a stool DNA multitarget assay panel for the detection of advanced colorectal neoplasia. *Clinical Colorectal Cancer*, 2003. 3(1): p. 47-53.

Traverso, G., A. Shuber, L. Olsson, B. Levin, C. Johnson, S.R. Hamilton, K. Boynton, K.W. Kinzler, and B. Vogelstein, Detection of proximal colorectal cancers through analysis of faecal DNA. *Lancet*, 2002. 359(9304): p. 403-4.

Trends in screening for colorectal cancer--United States, 1997 and 1999. *MMWR Mortal Wkly Rep*, 2001. 50: p. 162-66.

Tsuda, S., B. Veress, E. Toth, and F.T. Fork, Flat and depressed colorectal tumours in a southern Swedish population: a prospective chromoendoscopic and histopathological study. *Gut*, 2002. 51(4): p. 550-5.

Verne, J.E., R. Aubrey, S.B. Love, I.C. Talbot, and J.M. Northover, Population based randomized study of uptake and yield of screening by flexible sigmoidoscopy compared with screening by faecal occult blood testing. *BMJ*, 1998. 317(7152): p. 182-5.

Villa, E., A. Dugani, A.M. Rebecchi, A. Vignoli, A. Grottola, P. Buttafoco, L. Losi, M. Perini, P. Trande, A. Merighi, R. Lerose, and F. Manenti, Identification of subjects at risk for colorectal carcinoma through a test based on K-ras determination in the stool. [comment]. *Gastroenterol*, 1996. 110(5): p. 1346-53.

Walsh, J.M. and J.P. Terdiman, Colorectal cancer screening: scientific review. [see comment]. *JAMA*, 2003. 289(10): p. 1288-96.

Walter, S.D., D.J. Frommer, and R.J. Cook, The estimation of sensitivity and specificity in colorectal cancer screening methods. *Cancer Detect Prev*, 1991. 15(6): p. 465-9.

Weaver, P., B. Harrison, G. Eskander, et al., Colon cancer in blacks. A disease with a worsening prognosis. *J Natl Med Assoc*, 1991. 83: p. 133-38.

Wherry, D.C. and W.M. Thomas, The yield of flexible fiberoptic sigmoidoscopy in the detection of asymptomatic colorectal neoplasia. [see comment]. *Surg Endosc*, 1994. 8(5): p. 393-5.

Williams, A.R., B.A. Balasooriya, and D.W. Day, Polyps and cancer of the large bowel: a necropsy study in Liverpool. *Gut*, 1982. 23(10): p. 835-42.

Williams, C.B., R.H. Hunt, H. Loose, R.H. Riddell, Y. Sakai, and E.T. Swarbrick, Colonoscopy in the management of colon polyps. *Br J Surg*, 1974. 61(9): p. 673-82.

Winawer, S.J. and M. Shike, Prevention and control of colorectal cancer, in *Cancer Prevention and Control*, P. Greenwald, B.S. Kramer, and D.L. Weed, Editors. 1995, Marcel Dekker: New York. p. 537-539.

Winawer, S.J., A.G. Zauber, M.N. Ho, et al., Prevention of colorectal cancer by colonoscopic polypectomy. *N Engl J Med*, 1993. 329: p. 1977-81.

Winawer, S.J., R.H. Fletcher, L. Miller, et al., Colorectal cancer screening: clinical guidelines and rationale. *Gastroenterol*, 1997. 112: p. 594-62.

Wong, B.C., W.M. Wong, J.K. Chan, K.C. Lai, W.H. Hu, C.K. Chan, S.K. Lam, and D.L. Carr-Locke, Virtual colonoscopy for the detection of colorectal polyps and cancers in a Chinese population. *J Gastroenterol Hepatol*, 2002. 17(12): p. 1323-7.

Wong, W.M., S.K. Lam, K.L. Cheung, T.S. Tong, P. Rozen, G.P. Young, K.W. Chu, J. Ho, W.L. Law, H.M. Tung, H.K. Choi, Y.M. Lee, K.C. Lai, W.H. Hu, C.K. Chan, M.F. Yuen, and B.C. Wong, Evaluation of an automated immunochemical fecal occult blood test for colorectal neoplasia detection in a Chinese population. *Cancer*, 2003. 97(10): p. 2420-4.

Yee, J., G.A. Akerkar, R.K. Hung, A.M. Steinauer-Gebauer, S.D. Wall, and K.R. McQuaid, Colorectal neoplasia: performance characteristics of CT colonography for detection in 300 patients. *Radiology*, 2001. 219(3): p. 685-92.

Ziegler, R.G., S.S. DeVesa, and J. Fraumeni, Epidemiologic patterns of colorectal cancer, in *Important Advances in Oncology*, V.T. Devita, V.T., S. Hellman, and S.A.Rosenberg, Editors. 1986, Philadelphia: J. B. Lippincott.

Appendix K

Cost Issues in Cost Effectiveness Modeling of Colorectal Cancer Screening
Martin L. Brown, Ph.D.

SLIDE 1

Appendix K
Cost Issues in Cost Effectiveness
Modeling of Colorectal Cancer
Screening

Martin L. Brown, Ph.D.
Health Services and Economics Branch
Applied Research Program
Division of Cancer Control and Population Sciences
National Cancer Institute

1

SLIDE 1 NOTES: No notes.

SLIDE 2

Overview

- Cost Domains

- Cost Continuum

- Cost Components

- Unrelated Future Costs

- Sources of Data

- Sources of Variation in Cost Assumptions   2

SLIDE 2 NOTES: No notes.

SLIDE 3

Cost Domains

- Direct Medical Costs

- Time Costs

- Productivity Costs

3

SLIDE 3 NOTES: Cost domains refers to the types of costs included in a CEA, such as direct medical costs, time costs and productivity costs. Time costs – the cost to the patient due to lost time of receiving health care–are especially important in large-scale screening programs where a lot of people spend a lot of time in order for a few to benefit from early detection or prevention of colorectal cancer. Productivity costs may be relevant in populations where individuals must take time away from work to get screened. Most models have ignored these two dimensions and focused on the direct medical care costs.

SLIDE 4

Cost Continuum

- Initial Screening
    – Including Cost of Adverse Events
- Diagnostic Follow-up
    – Including Follow-up Biopsy
- Surveillance
- Treatment
    – Initial
    – Continuing
    – Terminal

4

SLIDE NOTES 4:   Across the cost continuum–from screening through follow-up, surveillance and treatment of cancer–there is great variability among models in what is assumed. In particular, accounting for the costs of treating colorectal cancer varies across models as to whether it is entered as a one-time package cost or whether it is entered over time and varies with the subsequent life of the patient. The latter approach acknowledges the high initial and terminal costs of treating colorectal costs, while the continuing care components recognizes that costs of treating and monitoring colorectal cancer continue over the course of a person's life following the initial treatment period.increased by about 2 years in a cohort of 50-year-old individuals. That 2-year difference leads to substantially lower life-expectancies in our model.

SLIDE 5

Cost Components of Screening

- Promotion and Recruitment

- Facilities, Equipment and other Infrastructure

- Time Costs

- Quality Control and Program Evaluation

5

SLIDE 5 NOTES: The costs listed above may all be relevant to a cost-effectiveness analysis of a screening program. However, the costs of promotion and recruitment have been largely ignored in most cost-effectiveness analysis. Those costs are basically fixed costs that deliver a certain level of participation in a screening program. These costs have been ignored in models focusing on the US population because in our health system, promotion and recruitment are typically not addressed through explicit organized and government-run screening programs. We have a mixture of publicity and promotion and advertising and public relations, as well as physician-led recruitment. We don't know much about how much of this activity occurs, much less about how much it costs.

In other high-income countries, where governments explicitly organize programs to screen, these fixed costs actually can be and do get estimated in the cost side of the cost-effectiveness analysis.

Facilities and equipment costs are implicitly included as amortized costs in procedure costs. (The Medicare fee structure, for example, was constructed with such amortized costs in mind.) But, when the fixed costs are incurred up front, in a short period of time, they may become financial constraints that health care providers or plans must take into account. Such issues are not considered in the models.

Time costs are typically not included in models, most likely because we simply do not have good estimates of such costs. NCI is currently conducting work that will provide estimates of time costs associated with treatment of CRC by phase of disease. It would be relatively straight-forward to get time costs for screening and diagnostic phases.

Quality control costs are typically not considered in the models of CRC screening. But such costs can be real, if not major. For example, a cost-benefit analysis of the Mammography Quality Standards Act estimated that quality control and program evaluation costs for that technology were on the order of $2 or $3 per examination, compared with total costs of $50 to $100 per examination. As colorectal cancer screening evolves in the US, explicit quality control programs will undoubtedly be required.

SLIDE 6

Sources of Cost Estimates

- Micro-costing at the program level
  - Done for some European trials, programs
- Time-motion studies
- Administrative Pricing
  - Medicare
  - Private Insurance (e.g. Medstat)
  - Single Institution Charges/Costs

6

SLIDE 6 NOTES: The first two sources of data for estimating procedure or program costs involve direct measurement of the inputs (and their values) that go into performing specific tasks involved in screening and treatment. They are not widely available, but they may merit more work, because the alternative source of costing data -- administered prices – are largely based on custom, history, and relative bargaining power of the buyers and sellers of procedures. They may be only loosely related to the true quantity and value of inputs required to produce the procedure or service. This is particularly true if the price of a screening procedure is based on its use as a diagnostic service, which can involve very different resource requirements. Consequently, I believe that relying on historical prices to represent the true resource cost of a mass screening program probably overestimates the true screening cost. Yet, it may be disingenuous to base a cost-

effectiveness analysis on an efficient price, drawn perhaps from a time-and-motion study, when in fact the patient or health plan will have to pay the price established through an administered pricing system. In that case, the realized cost may be two or three times higher than those assumed in the cost-effectiveness analysis.

SLIDE 7

Sources of Variation in Costs Across Studies

- Screening
  - Unit Costs
  - Procedures or Services that get counted
- Treatment
  - Initial vs. Long-term Costs
  - Accounting for Unrelated Future Costs
- Surveillance
  - Practice pattern assumed

7

SLIDE 7 NOTES: Here are some of the obvious sources of variation in estimated costs across the CEA studies we are examining in this workshop.

One very important issue is whether the study includes a charge for the additional unrelated future health care costs that would be incurred simply as a result of screened patients living longer. Later slides will refer to this issue in more detail.

SLIDE 8

Variations in Estimates of Sigmoidoscopy Unit Costs

Author	Year	Type	Cost
Geul	1997	Audit	$102
Norum	1998	Charge	$147
Cromwell	1998	Charge	$287
Lewis	1999	Audit	$166
Bolin	1999	Charge	$281
Wallace	1999	Audit	$298
Frazier	2000	Audit	$279
Khandker	2000	Claims	$349
Sonnenberg	2000	Claims	$406
Helm	2000	Claims	$465
Vijan	2001	Charge	$242
Suleiman	2001	Charge	$386
Lewis	2002	Charge	$215
Whynes	2003	Audit	$91

Adapted from Whynes, et al. 2003

8

SLIDE 8 NOTES: Here is a table showing the wide range of estimates of the unit cost of sigmoidoscopy among recent studies (Whynes et al., 2003). It is no wonder that we get differences in cost-effectiveness ratios when there is such a wide range of costs of the procedure. The screening cost accounts for a large proportion of the entire lifetime cost. A difference in costs by a factor of four or five would obviously affect the conclusions of an analysis.

SLIDE 9

Variations in Costs
Related to Components Considered

- Examples
    - Time and Travel Costs
    - Provider Costs
    - Biopsy Practice

9

SLIDE 9 NOTES: No notes.

SLIDE 10

Time and Travel Costs

- Cost of Sigmoidoscopy to NHS:
  – $90.50

- Cost of time and travel to individual screened:
  – $34.50

- Total cost of Sigmoidoscopy examination:
  – $125.00

Source: Whynes et al., 2003

10

SLIDE 10 NOTES: Here is an example of unit costs estimated in the context of a clinical trial of one-time sigmoidoscopic screening in England (Whynes et al., 2003). Note that the time cost as estimated in the trial is not a trivial proportion of the entire unit cost of screening.

SLIDE 11

Provider Costs

- Cost of FOBT assuming no provider contact:
  - $4.50

- Cost of FOBT assuming primary care visit:
  - $39.00

- What would be the cost of population-based recruitment?

Source: Frazier et al. 2000   11

SLIDE 11 NOTES: Here is an example of issues in costing of FOBT (Frazier et al., 2000). The current reimbursement from Medicare for an FOBT is about $4.50. Many studies assume that Medicare's allowed amount is the cost for FOBT.

However, if one must visit a primary care physician to receive and be instructed in the proper application of the test, the cost would be an additional $40. Even if you assume that the FOBT is only one-half the cost of the visit, the procedure cost increases dramatically.

Of course the question is whether that is a reasonable way to do it. We rely on physicians to recruit their patients, but we could also envision an FOBT screening program where recruitment is population-based and does not involve a physician encounter. (Many European FOBT screening programs follow that paradigm.) Appropriate costing of that screening paradigm would start with the $4.50 cost, and then add, say a fixed cost of $500,000 (for recruitment operations) divided over the number of individuals who are screened by such a program.

SLIDE 12

Biopsy Practice

- If a biopsy is not conducted on a suspicious lesions found by Sigmoidoscopy, a costly referral colonoscopy is triggered.

- A national survey conducted in 2000 found that about two-thirds of primary care physicians who perform screening sigmoidoscopy refer to gastroenterologists to conduct biopsy procedures for lesions found by sigmoidoscopy.

Source: Klabunde et al., 2003   12

SLIDE 12 NOTES: Here is an example of how a seemingly small detail of a screening protocol can have major implications for costs (Klabunde et al., 2003). It makes a big difference whether a lesion found on sigmoidoscopy is biopsied on the spot, or whether the patient is referred for colonoscopy. In a 2000 survey of primary care physicians who do flexible sigmoidoscopy (NOT gastroenterologists), we found that less than one-third said they would actually take a biopsy as part of the screening procedure. In practice, then, the majority of primary care practitioners do refer to colonoscopy, with major consequences for costs.

SLIDE 13

Variations in Treatment Cost Due to Future Unrelated Costs

- For cases of colorectal cancers that are prevented should the treatment cost related to extended life expectancy and competing causes of death be counted?

- A first approximation to taking future unrelated costs into account is to only count initial phase treatment costs.

13

SLIDE 13 NOTES: The question of how unrelated costs are treated is very important, because the cost-savings due to avoided treatment is a very large part of the net cost of a screening program. So, even modest changes in the estimate of cost savings can make huge differences in the final cost-effectiveness ratio.

We can ask, however, what would happen to a person who, because of screening, is spared a CRC. That person will eventually die of some other cause, and the cost of treating that condition may be similar to the cost of treating colorectal cancer. In most cases, we hope, it would be later in time, and when costs are discounted to their present value, there would be some net savings even if the absolute costs did not decline. In addition, during the extra years of life that a person lives, he or she would obtain medical care, which would also decrease the net savings from prevention of CRC.

There is no consensus, however, about whether the extra costs associated with longer life or competing causes of death should be included in CEA (Gold et al., 1996).

SLIDE 14

Cancer-Related Treatment Cost of Colorectal Cancer

Stage	Without Unrelated Costs	With Unrelated Costs
In Situ	$28,000	$26,000
Stage 1	$32,700	$29,000
Stage 2	$34,400	$22,700
Stage 3	$41,600	$17,300
Stage 4	$29,400	[$9,000]

Source: Etzioni et al. 2001

14

SLIDE 14 NOTES: What are the implications of including or excluding unrelated health care costs from the treatment cost estimate?

The table in this slide shows the results of a study in which we examined that question (Etzioni et al., 2001). We examined the treatment cost of colorectal cancer by stage. All of the costs in the first column are cancer related. However, if one includes the unrelated costs of living longer, the costs available to be averted through screening are reduced. The later the stage at diagnosis, the lower the potential cost savings from prevention become, and at the latest stage prevention actually increases costs. In that case, instead of saving costs of treatment, it actually costs $9,000 more. That is because individuals diagnosed in stage 4 tend to die quickly, and all the usual health care expenses are avoided.

I do not know what effect such different cost assumptions would have on the outcomes of CEAs, as it would depend upon the distribution of incidence across the stages of disease and the impact of a screening strategy on each particular stage. But the differences in cost imply that the results could be quite sensitive to the inclusion or exclusion of these costs.

SLIDE 15

Variations on Cost Due to Surveillance Practices

- What assumption should be made about the use of surveillance procedures following detection of a colonic adenoma?

15

SLIDE 15 NOTES: Surveillance assumptions are a large component of the total lifetime costs of a screening program. If the frequency of surveillance is low, or if it is conditioned on the size of the polyp or other criteria, the overall program costs will be lower than it would be if one were to assume a uniform policy of surveillance with colonoscopy every three years.

A related question in this regard is whether the assumptions regarding surveillance should reflect a 'best practices" or guidelines-based surveillance regimen, or a surveillance regimen that reflects actual practice.

SLIDE 16

Effects of Surveillance on Screening Program Costs

- Analysis of FOBT screening for the Medical population over a 30 year period
  - Surveillance according current guidelines – net program cost = $205 million per million participants
  - Surveillance once every three years – net program cost = $453 million per million participants

Source: van Ballegooijen et al., 2003.   16

SLIDE 16 NOTES: We addressed the question of the impact of using guidelines-based surveillance versus a strategy that assumes surveillance occurs ever 3 years (van Ballegooijen et al., 2003). That study was for an immunochemical FOBT test in a program that extended over 30 years and discounted at 3 percent per year. Total program costs almost doubled between current guidelines for surveillance and a uniform once-every-three-years surveillance practice.

SLIDE 17

Conclusion

- Differing approaches can result in major differences in cost assumptions across models.

17

SLIDE 17 NOTES: No notes.

SLIDE 18

Additional Considerations

- Are unit cost estimates based on historic diagnostic practice appropriate for population-based screening?
- Should delivery organizational setting be considered?
- Should time-horizon be considered in relationship to depreciation of sunk investment?

18

SLIDE 18 NOTES: There is a paucity of studies that try to establish the most efficient cost of providing screening services. That is, how should such programs be organized and delivered so as to deliver the particular services associated with a screening strategy in the most efficient manner.

For example, assuming that flexible sigmoidoscopy screening would be carried out mainly by primary care physicians, the cost implications would be different from those under a sigmoidoscopy screening program carried out by gastroenterologists, or by screening endoscopy mills, a term I use in a positive sense, because of their potential to deliver endoscopic services efficiently.

Today, we have almost no information on the relative costs of delivering screening services in those different organizational environments. I am not even sure where such a responsibility for sponsoring such studies would lie, whether at NIH or at AHRQ.

REFERENCES

Etzioni R, Ramsey SD, Berry K, Brown M. 2001. The impact of including future medical care costs when estimating the costs attributable to a disease: A colorectal cancer case study. *Health Economics.* 10(3): 245–256.

Frazier AL, Colditz GA, Fuchs CS, Kuntz KM. 2000. Cost-effectiveness of screening for colorectal cancer in the general population. *JAMA.* 284(15): 1954–1961.

Gold MR, Siebel JE, Russell LB, Weinstein MC, Editors. 1996. *Cost-Effectiveness in Health and Medicine.* New York: Oxford University Press.

Klabunde CN, Frame PS, Meadow A, Jones E, Nadel M, Vernon SW. 2003. A national survey of primary care physicians' colorectal cancer screening recommendations and practices. *Prev Med.* 36(3): 352–362.

Mandel JS, Church TR, Ederer F, Bond JH. 1999. Colorectal cancer mortality: Effectiveness of biennial screening for fecal occult blood. *J Natl Cancer Inst.* 91(5): 434–437.

van Ballegooijen M, Habbema JDF, Boer R, Zauber AG, Brown ML. 2003. Report to the Agency for Health Care Research and Quality: A Comparison of the Cost-Effectiveness of Fecal Occult Blood Tests with Different Test Characteristics in the Context of Annual Screening in the Medicare Population.

Whynes DK, Frew EJ, Edwards R, Atkin WS. 2003. Costs of flexible sigmoidoscopy screening for colorectal cancer in the United Kingdom. *Int J Technol Assess Health Care.* 19(2): 384–395.

The effect of fecal occult-blood screening on the incidence of colorectal cancer. *N Engl J Med.* 343(22): 1603–1607.

Vijan S, Hwang EW, Hofer TP, Hayward RA. 2001. Which colon cancer screening test? A comparison of costs, effectiveness, and compliance. *Am J Med.* 111(8): 593–601.

Appendix L

Current Evidence on Compliance
Sally W. Vernon, M.A., Ph.D.

SLIDE 1

Appendix L
Current Evidence on Compliance

**IOM Workshop on Economic Models of Colorectal Cancer (CRC)
Screening in Average-Risk Adults**

**Sally W. Vernon, M.A., Ph.D.
University of Texas-Houston
School of Public Health
January 26-27, 2004**

1

SLIDE 1 NOTES: No notes

208

SLIDE 2

SLIDE 2 NOTES: I was asked to answer, or try to answer, 5 questions related to compliance with colorectal cancer (CRC) screening. This presentation represents an update on reviews I have conducted with colleagues in the recent past (Peterson and Vernon, 2000; Vernon, 1997). The answers to some of the questions, such as the first two, are straightforward, and we have a reasonable amount of published data to draw on in answering those questions.

The 3rd question requires some extrapolation from existing data because it hasn't been studied directly. I will share some thoughts about how we might arrive at a best guess, but I think this question warrants discussion.

There aren't much published data to address the 4th question for any cancer screening test. But I will summarize the available data.

And to answer the 5th question, I plan to turn the podium over to my colleague Dr. Seeff, who has led a study by the CDC that may provide insight to help us answer this important question.

SLIDE 3
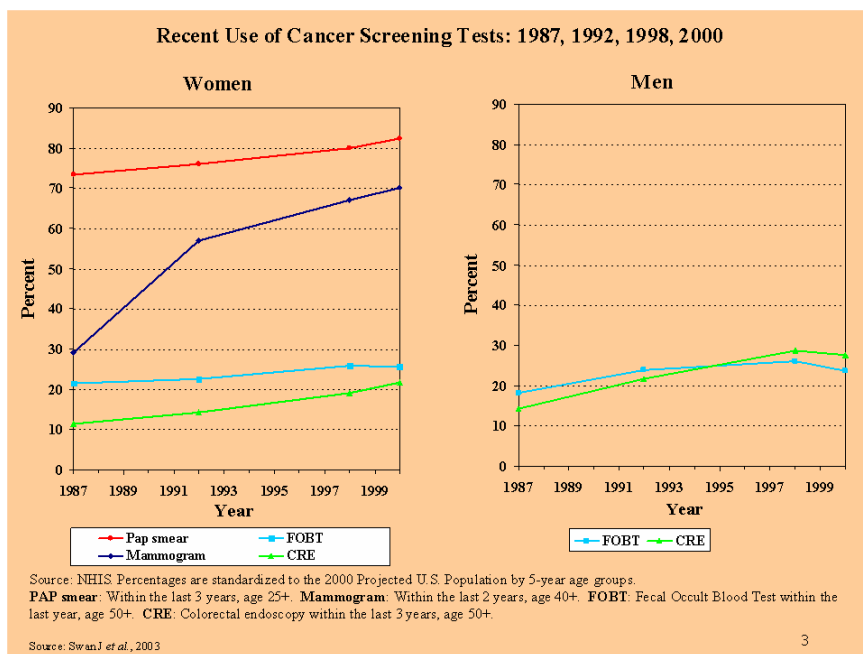
SLIDE 3 NOTES: What is known about the levels of, and patterns over time in, participation in CRC screening in the U.S. population?

There are 2 national data sources that have collected self-report data on use of CRC screening tests – the National Health Interview Survey (NHIS) and the Behavioral Risk Factor Surveillance System (BRFSS). These two surveys use different data collection methods. The NHIS is a face-to-face interview, and the BRFSS is a telephone interview. There also are differences in the question wording and the time intervals asked about, both within and between surveys. Another survey conducted by the NCI called the Health Information and Trends Survey (HINTS) was completed by the NCI and the data are forthcoming. So we are beginning to get a more complete and up-to-date picture of CRC screening test use.

Data on CRC test use have been collected by the NHIS since 1987. The data show that, compared with other types of cancer screening tests, tests for colorectal cancer are infrequently used among both men and women, and that their use is increasing only gradually. In 2000, less than 30% of the population had had a fecal occult blood test or endoscopy within the time interval recommended by the American Cancer Society.

At present, the available survey data do not permit a detailed look at the patterns of screening over time and across individuals. For example, whether a 60 percent screening rate in the population in each year represents the same individuals over time, or different

individuals, is critical to assessing the impact of compliance on screening costs and effectiveness. One survey currently in analysis at NCI may give some hints, because it asks individuals not only about the most recent screening encounter, but also about the next most recent. And the BFRSS may ask more refined questions about such patterns in future surveys.

SLIDE 4

Trends in Recent* Fecal Occult Blood Test Prevalence (%), by Educational Attainment and Health Insurance Status, Adults 50 Years and Older, U.S., 1997-2002
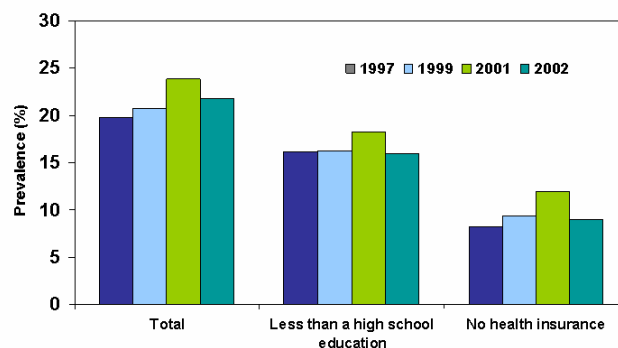
*A fecal occult blood test within the past year. Note: Data from participating states and the District of Columbia were aggregated to represent the United States.
Source: Behavioral Risk Factor Surveillance System CD-ROM (1996-1997, 1999) and Public Use Data Tape (2001, 2002), National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention and Prevention, 1999, 2000, 2002, 2003.

SLIDE NOTES 4: Data from the other national survey, the BRFSS, cover the time period 1997 to 2002. Despite an increase in FOBT use among both men and women from 1999 to 2002, prevalence remains low.

This slide also includes data that, in part, address the 2nd question: "What factors affect adherence to different kinds of technologies?"

Adults with less than a high school education are less likely to report a recent FOBT. And the prevalence for adults with no health insurance is approximately 10 percentage points lower than the prevalence for all adults – less than 10% in every years except 2001.

SLIDE 5

Trends in Recent* Sigmoidoscopy or Colonoscopy
Prevalence (%), by Educational Attainment and Health
Insurance Status, Adults 50 Years and Older, U.S., 1997-
2002
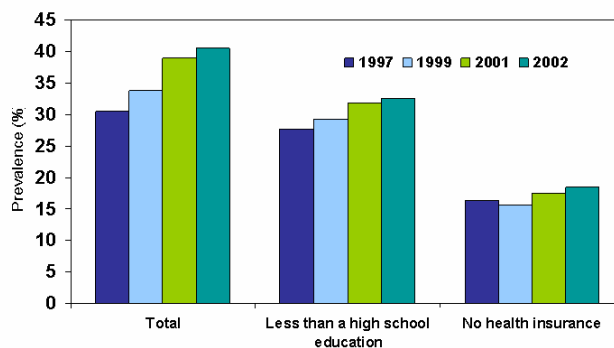
*A flexible sigmoidoscopy or colonoscopy within the past five years. Note: Data from participating states and
the District of Columbia were aggregated to represent the United States.
Source: Behavioral Risk Factor Surveillance System CD-ROM (1996-1997, 1999) and Public Use Data Tape
(2001, 2002), National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease
Control and Prevention and Prevention, 1999, 2000, 2002, 2003.

5

SLIDE 5 NOTES: The prevalence of sigmoidoscopy (SIG) or colonoscopy within the past five years, while low, increased from 30% to 40% between 1997 and 2002. The prevalence and the rate of increase are higher than for recent FOBT.

As with FOBT, adults with less than a high school education were less likely to report SIG or colonoscopy than all adults. And the prevalence for adults with no health insurance is approximately 15 percentage points lower than the prevalence for all adults, and is less than 20% in all years.

SLIDE 6

Factors Positively Associated with
Colorectal Cancer Screening: NHIS 2000

- Age $\geq$ 65
- White ethnicity
- More formal education
- Higher income
- **Having health care coverage**
- **Having a usual source of care**
- **Frequency of medical care visits**

6

SLIDE 6 NOTES: This slide summarizes the results of research on factors that are associated with having a CRCS test. Factors are similar for FOBT, SIG, and COL. One of the messages here, and in the prior 2 slides, is that access to medical care -- such as having health insurance, a healthcare provider, or frequent contact with the medical care system -- is an important factor in increasing the chances of being screened for CRC.

These factors are similar to those for getting mammography screening and Pap testing.

Another factor that has not been looked at in national surveys but that is strongly associated with adherence to cancer screening recommendations, including CRC screening, is having a physician recommend the test. Data from the National Cancer Institute's Health Information and Trends Survey (HINTS) of 2003 are currently being analyzed, but it appears that a small proportion of patients report having their physicians recommend colorectal cancer screening. So, there is room for improvement on that score.

SLIDE 7

Trends in Recent* Mammogram and Pap Test Prevalence (%), U.S., 1991-2002
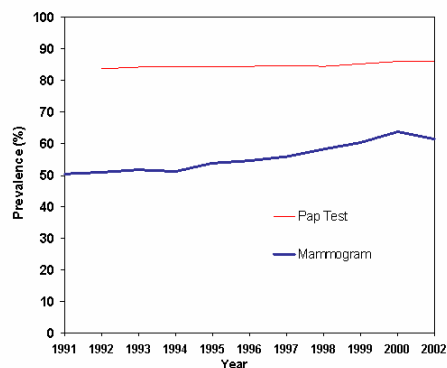
•A mammogram in the past year, women 40 and older. A Pap test within the past 3 years, women 18 and older.
•Source: Behavior Risk Factor Surveillance System CD-ROM (1984-1995, 1996-1997, 1998, 1999) and Public Use Data
Tape (2000, 2002), National Centers for Chronic Disease Prevention and Health Promotion, Centers for Disease Control
and Prevention 1997, 1999, 2000, 2000, 2001,2003.

SLIDE 7 NOTES: What is our "best guess" about the level of compliance that might be expected with new screening technologies such as DNA and virtual colonoscopy?

I think that the best way to address this question is by looking at the prevalence and trends in prevalence for the use of other cancer screening tests and associated technologies.

Of all cancer screening tests for which we have data, Pap testing prevalence is the highest. Rates have been steady at ~85% since 1992. The prevalence of women reporting a recent mammogram has increased steadily since 1991 and was 62% in 2002.

SLIDE 8

Recent* Prostate-Specific Antigen (PSA) Test Prevalence (%) and Digital Rectal Exam Prevalence (%), Men 50 Years and Older, U.S., 2001-2002
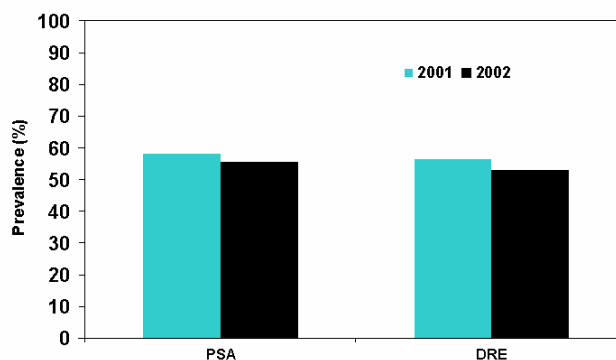
*PSA test or DRE test within the past year. Note: Data from participating states and the District of Columbia were aggregated to represent the United States.
Source: Behavioral Risk Factor Surveillance System Public Use Data Tape (2001, 2002), National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention, 2002, 2003.

SLIDE 8 NOTES: Although PSA is not a recommended for population-based screening, the data are of interest because the technology used. In 2002, the percentage of men who had a PSA test within the past year was 55% while the percentage of men who had a DRE within the past year was 53%.

The question I would pose from the data viewed collectively is: What makes Pap testing different from the other tests, particularly DRE and PSA? They all are relatively inexpensive, require contact with a physician, can be done easily in the context of a primary care visit, and they aren't very different in terms of their invasiveness. One obvious difference is that the data on Pap testing are for women while the data from DRE and PSA are for men. We know that women are more likely to have contact with the medical care system than men.

Another possible reason is that PSA is not a test that is recommended by the US Preventive Services Task force for population-based screening. Physicians have been encouraged to discuss PSA testing with their patients to allow an informed decision about whether to engage in it. But, from my own ongoing research, it appears that physicians are not generally engaging in such discussions. So, I believe it may not be done much because it has not disseminated into practice as a routine test.

Another observation from these data is that rates of recent mammography aren't as high as we would like from a public health perspective – they are around 60%. Mammography is a more inconvenient procedure than DRE, PSA or Pap testing because it typically requires a referral which often means coming back for another appointment. While not invasive, it's more uncomfortable than DRE, PSA and Pap tests. And it's potentially more expensive.

What, then, is reasonable to expect for CRC screening tests and procedures? FOBT is similar in some respects to DRE, PSA, and Pap testing in that it is relatively inexpensive. It is different in that it may or may not require contact with a physician to acquire a test kit. SIG and COL are similar to mammography in that they typically require a referral, require more planning, and are arguably even more uncomfortable or unpleasant than mammography, although they are more invasive.

If we had a test that was inexpensive, easy for physicians to do, and acceptable to patients, what might be possible? I would argue that something close to 80% is possible. If we don't have such a test for awhile, what might be some other issues that affect CRC screening prevalence, particularly for colonoscopy?

SLIDE 9

Issues Affecting Adherence to Colonoscopy

- **Greater importance of "system" issues**
  - lack of insurance coverage
  - capacity
- **Patient issues**
  - patient confusion over "best" test
  - time lost from work
  - transportation assistance required

9

SLIDE 9 NOTES: There are also issues associated with physicians that are most interesting. Work by Dr. Klabunde at NCI is currently conducting show that physicians' perceptions of their recommending tests or performing tests on patients are exceedingly high (Klabunde et al., 2003). Roughly 60 to 70 percent of physicians say that they are recommending or performing colon cancer screening tests on their average-risk patients. Those are not consistent, as you can see, with the prevalence estimates for these tests. So, there is some disconnect that merits more exploration.

SLIDE 10

What is known about compliance with follow-up for a positive test and for post-polypectomy surveillance?

- 45 observational studies of follow-up
  - 11 of abnormal FOBT or SIG

- Definitions of follow-up varied widely from consultation to completion of all recommended diagnostic tests

- Measurement varied across studies
  - Measurement method affected prevalence estimates even within the same study

- In ~66% of all studies of cancer screening tests, follow-up was below 75%

10

SLIDE 10 NOTES: All studies were regional; there are no national data on follow-up. The majority of studies reviewed by Dr. Yabroff and her colleagues were conducted in clinics, programs, or health care systems that had mechanisms for tracking follow-up care (Yabroff et al., 2003; Bastani et al., under review). The actual prevalence of follow-up may be lower outside of organized systems of care or for the uninsured.

A study by Myers and colleagues (Myers et al., 2001) compared measurement of complete diagnostic evaluation after a positive FOBT by 4 methods – external chart review, internal chart audit, administrative data review, and a combination of chart review and administrative data notes – and found that rates ranged from 49 to 79%. Regardless of how measured, prevalence is less than optimal and may compromise the effectiveness of mass screening programs. We do not know the impact of delayed, incomplete, or no follow-up on stage of disease at diagnosis because population-based data on follow-up care are not routinely collected.

Dr. Yabroff and her colleagues conclude that: "Increasing the availability of screening follow-up data will be an important area for future evaluations of the quality of care provided in screening programs."

To my knowledge, there are no published data of post-polypectomy surveillance.

SLIDE 11

How do the organization of and
geographic availability of screening
services affect the level of compliance
that can be expected?

11

SLIDE 11 NOTES: This is a compelling question that needs more research. At present, there is little evidence available to report.

REFERENCES

Bastani R, Yabroff KR, Myers R, Glenn B. Interventions to improve follow-up of abnormal
        findings in cancer screening. *Cancer*, under review.
Hawley ST, Vernon SW, Levin B, Vallejo B. Prevalence of colorectal cancer screening in a
        large medical organization. *Cancer Epidemiol Biomarkers Prev*, in press.
Klabunde CN, Frame PS, Meadow A, Jones E, Nadel M, Vernon SW. 2003. A national survey
        of primary care physicians' colorectal cancer screening recommendations and practices.
        Prev Med. 36(3): 352–362.
Myers RE, et al. 2001. Measuring complete diagnostic evaluation in colorectal cancer screening.
        Cancer Detect Prev. 25(2): 174–82.
Peterson SK, Vernon SW. 2000. A review of patient and physician adherence to colorectal
        cancer screening guidelines. *Seminars in Colon & Rectal Surgery*. 11: 58–72.

Swan J, Breen N, Coates RJ, Rimer BK, Lee NC . 2003. Progress in cancer screening practices in the United States: Results from the 2000 National Health Interview Survey. *Cancer. Vol. 97*(6): 1528-1540.

Vernon SW. 1997. Participation in colorectal cancer screening: a review. *J Natl Cancer Inst.* 89(19): 1406–1422.

Yabroff KR, Washington KS, Jacobs A, Neilson E, Mandelblatt JS. 2003. Is the promise of cancer screening programs being compromised? Quality of follow-up care after abnormal screening results. *Med Care Res Rev.* 60(3): 294–331.

Appendix M

Recent Trends in Follow-up
Surveillance in Medicare Beneficiaries
Todd Anderson, M.S.

SLIDE 1

Appendix M
Recent Trends in Follow-up and
Surveillance in Medicare Beneficiaries

Todd Anderson
Congressional Budget Office
January 26, 2004

Note: The views expressed in this talk are those of the
author and should not be interpreted as those of
the Congressional Budget Office.

1

SLIDE 1 NOTES: The information presented here is based on an analysis of a
longitudinal Medicare data set. Please note that the data presented here do not reflect the
views of CBO.

SLIDE 2

Questions

1. Frequency of colonoscopy in the years following index polypectomy.

2. Patterns of short-term follow-up following sigmoidoscopy.

2

SLIDE 2 NOTES: My talk reviews two issues: (1) How frequently do patients who have undergone a polypectomy receive surveillance colonoscopies in subsequent years? And (2) What kinds of follow-up procedures occur in patients, patients with polyp(s) discovered on a a sigmoidoscopy?

SLIDE 3

Data

1.   Longitudinal 5% random sample of Medicare claims data, beneficiary level and monthly, from 1989-1999.

2.   Newly eligible beneficiaries are added to sample each month to maintain representative panel in each year.

3.   Fee-for-service Beneficiaries only, 10 to 15 percent of beneficiaries enrolled in managed care are excluded.

4.   Final Sample Size = 3.4 million unique individuals (1.6-1.7 million beneficiaries per year).

3

SLIDE 3 NOTES: The longitudinal data base currently spans an 11-year history of claims for a sample of Medicare beneficiaries.

The sample is updated in every month, with the addition of a sample of only newly eligible beneficiaries and the deletion of those who die.

The particular subsample on which this analysis is based is for people who were continuously enrolled in fee-for-service (i.e., traditional) Medicare. About 10-15 percent of Medicare beneficiaries enrolled in Medicare managed care plans in any year, and we have no claims data for that group.

Claims data in this dataset are grouped by month. Therefore, we can only know whether an individual had at least one service of a particular type (e.g., sigmoidoscopy) but cannot determine the exact date(s) on which such a service occurred.

SLIDE 4

Colonoscopy frequency following a Polypectomy:
Methodology

- Starting in 1994 (but allowing for later entry for new enrollees), each beneficiary is included in sample until:
  - -- Death
  - -- Initial disenrollment from FFS, or
  - -- End of 1999 (sample size=2.4 million).

- Selected all who had a colonoscopy or sigmoidoscopy with polypectomy (sample size=125,000) -- *Upper Bound*

- Identified month of first colonoscopy at least 6 months after polypectomy

- Dropped those with any cancer-related diagnosis code in two different months between 1991 and 1999 -- excluding about 65,000 beneficiaries -- *Lower Bound*

4

SLIDE NOTES 4: To address the first question: What happens to patients in the years following a polypectomy with respect to receipt of a colonoscopy? The sample consisted of individuals present in 1994 and later years. Each individual was followed until death, disenrollment from fee-for-service Medicare; or the end of the study period (1999).

Of that group, any individual who had undergone a polypectomy, either by sigmoidoscopy or full colonoscopy was identified.

I then identified the month of the first colonoscopy that occurred at least 6 months following the polypectomy. Colonoscopies performed within 6 months of the index polypectomy could have been follow-up procedures related to the polypectomy. We were interested in long-term surveillance procedures.

Finally, we tested the effect on the analysis of eliminating any individual with any kind of cancer diagnosis. Because we were dealing with claims data, we had to rely on a record of an ICD-9 cancer diagnosis in any claim over the beneficiary's claim history. We defined as "potential cancer beneficiaries" any individuals who had a cancer diagnosis in at least two different months in the period.

SLIDE 5

Percent of Beneficiaries with a Colonoscopy Following a Polypectomy by Years of Enrollment After Initial Service (N=125,000)
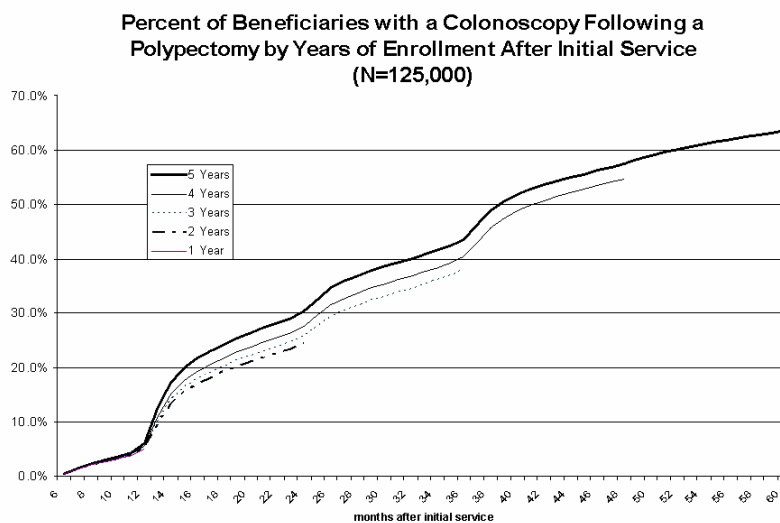
SLIDE 5 NOTES: This chart describes the percent of beneficiaries who received their first post-polypectomy colonoscopy in or before the month shown on the x-axis. Because the sample included individuals with different follow-up periods, we provide separate estimates for individuals at least 1 (but not 2) years of follow-up; at least 2 years of followup; and so on, until the maximum of 5 years of followup.

There are modest differences in the cumulative frequency curves for each of the groups. Those with shorter follow-up periods tended to have slightly lower rates of colonoscopy than those with longer periods.

In examining data among various demographic subgroups, I noticed some interesting trends among some smaller subgroups. However, the number of individuals sampled in those groups was too small to draw any significant conclusions. The probability of follow up among patients by race, sex, age and income may be worth closer examination with larger databases.

SLIDE 6

Percent of Non-Cancer Beneficiaries with a
Colonoscopy Following a Polypectomy
by Years of Enrollment After Initial Service (N=60,000)
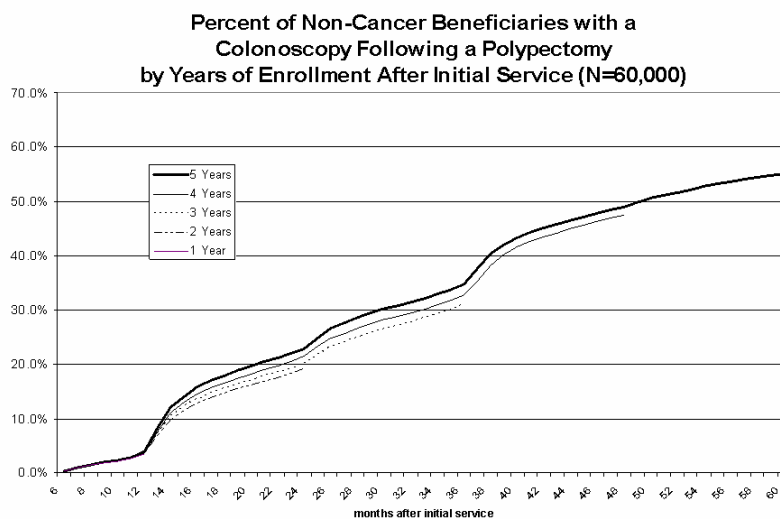
SLIDE 6 NOTES: This chart shows the same results as the previous group, except that this is the sample that excludes possible cancer beneficiaries. The pattern is similar as for the larger group, but this group is somewhat less likely to have a subsequent colonoscopy. For those with 5 years of data available, 55 percent had had at least one colonoscopy in the period between 6 months and 5 years following the index polypectomy.

SLIDE 7

Limitations

- Analytic treatment of potential cancer beneficiaries too undifferentiated
- Cannot distinguish among types or size of polyps removed
- Sample attrition correlated with health status
- Can't differentiate between screening and non-screening triggering events.

7

SLIDE 7 NOTES: The results shown up to this point have several important limitations. Because the database is exclusively Medicare claims for payment, it is impossible to differentiate among types of polyps removed, or sizes. We also cannot differentiate between polypectomies that arose out of a screening examination and those that occurred for diagnostic reasons. Nor can we differentiate between subsequent colonoscopies done for surveillance purposes and those done for diagnostic purposes.

SLIDE 8

Follow-up after sigmoidoscopy:
Methodology

- From individuals in 1993-1999 sample, selected those with fewer than two months in which a Cancer diagnosis was assigned between 1991 and 1999
- Defined **triggering event** as a sigmoidoscopy in an individual with at least 12 months of FFS enrollment history <u>AND</u> no sigmoidoscopy or colonoscopy in the previous six months
- Further excluded individuals with barium enema in previous six months <u>OR</u> in month of triggering event, or those without six months of FFS enrollment after the month of the triggering event
- Sample size = 83,198 sigmoidoscopies
- Identified all polypectomies in the month of the triggering event or in subsequent six months (6,257 cases). Further identified sigmoidoscopies with polypectomy in the initial service (1,218 cases).

8

SLIDE 8 NOTES: Now for the second question: What happens to people who receive a sigmoidoscopy in the months immediately following the procedure? In particular, how are positive sigmoidoscopy examinations followed up?

In this case we examined individuals in the sample in the period 1993-1999 and we eliminated all individuals who met the "possible cancer beneficiary" criterion.

We defined a triggering event as any sigmoidoscopy that met the following criteria:
- performed on a patient with at least 12 months of history in the sample,
- no history of a sigmoidoscopy or colonoscopy in the previous 6 months;
- no barium enema in the same or previous six months;
- at least six months' worth of data available after the index date.

Those criteria resulted in 83,000 index (triggering) events. We zeroed in on 6,257 of those with polypectomies in the month of the triggering event. And we further scrutinized about 1,200 beneficiaries whose index sigmoidoscopy included the removal of a polyp.

SLIDE 9

**Type of Polypectomy During the Six Months
Following Any Sigmoidoscopy (N=6,257)**
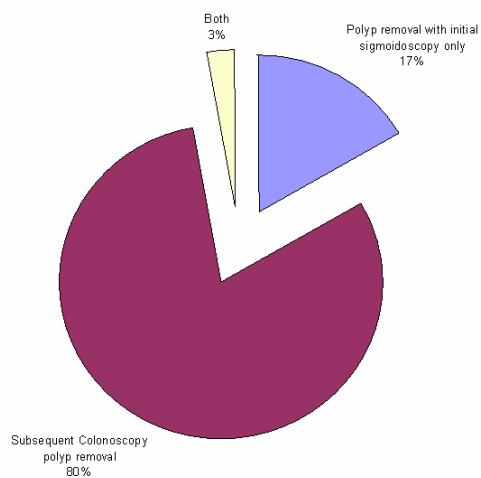
SLIDE 9 NOTES: In the universe of 6,257 sigmoidoscopy examinations associated with a polypectomy at any point in the succeeding six-months, fully 80 percent had the polypectomy in a subsequent colonoscopy, not in the initial sigmoidoscopy.

Seventeen percent of the polypectomies occurred as part of the triggering sigmoidoscopy examination.
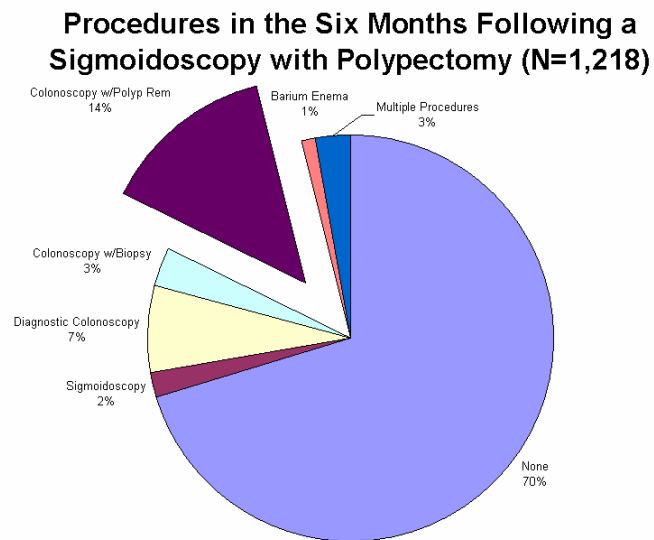
SLIDE 10

SLIDE 10 NOTES: This chart examines in greater detail what happened to the 17 percent of polypectomies that occurred as part of the sigmoidoscopy, in the months following sigmoidoscopy. Here, 25 percent of the cases went on to have at least one other colorectal diagnostic procedure in the six months following the triggering event.

SLIDE 11

Limitations

- Cannot differentiate among types or sizes of polyps removed
- Cannot differentiate between screening and non-screening triggering events

11

SLIDE 11 NOTES: The limitations of this part of the analysis are similar to those we encounter with the earlier analysis.

SLIDE 12

Key Findings

- 55% to 64% of Medicare beneficiaries having a colonoscopy or sigmoidoscopy with polypectomy obtained at least one surveillance colonoscopy in a 5-year period

- A large majority of Medicare beneficiaries who underwent polypectomy within six months of a sigmoidoscopy had polyps removed in subsequent procedures

12

SLIDE 12 NOTES: We conclude that somewhere between 55 and 64 percent of Medicare beneficiaries who undergo a polypectomy have at least one subsequent colonoscopy in the 5 year surveillance window.

   Diagnostic and therapeutic follow-up following a sigmoidoscopy is frequent, and a large majority of those how underwent polypectomy within six months of the sigmoidoscopy had their polyps removed in subsequent procedures.

Appendix N

Preliminary Results from CDC's Estimate of the National Capacity for Colorectal Cancer Screening and Follow-Up
Laura C. Seeff, M.D.

SLIDE 1

Appendix N
Preliminary Results from CDC's Estimate of the
National Capacity for Colorectal Cancer
Screening and Follow-Up

Laura C. Seeff M.D.
Division of Cancer Prevention and Control
Centers for Disease Control and Prevention,
Atlanta Georgia

1

SLIDE 1 NOTES: In this presentation, I will briefly show preliminary results of CDC's capacity assessment. I will touch on the methods and results of the study listed above. We hope to have the final results published in the late Spring (Seeff et al., 2004).

SLIDE 2

Capacity Assessment Overview

- National Survey of Endoscopic Capacity (SECAP)
  - Estimate current volume of lower endoscopic procedures

- Forecasting model
  - Estimate unmet need for lower endoscopies

- Comparison of volume to unmet need
  - Capacity assessment

2

SLIDE 2 NOTES: We assessed the national capacity to perform endoscopic procedures in three steps. We first implemented the National Survey of Endoscopic Capacity, or SECAP, to estimate the current volume of lower endoscopic procedures currently being performed. We then designed a forecasting model to estimate the unmet need for a lower endoscopy. Finally, we compared the two to make our capacity assessment.

SLIDE 3

SECAP Methods

- Sampling frame based on endoscopy sales to medical practices, 1996-2000
- 1809 surveys sent; 1346 (74.5%) returned
- Current and maximum potential weekly volume of flexible sigmoidoscopy and colonoscopy

3

SLIDE 3 NOTES: We used a sampling frame based on endoscope sales made to medical practices between 1996 and 2000.

We asked survey recipients, among many other questions, what is the current and maximum potential weekly volume of both flexible sigmoidoscopy and colonoscopy, given the current resources without making any additional investments or changes? "Current resources refers" to personnel, though in other parts of the survey we asked about equipment.

SLIDE 4

Annual Lower Endoscopies, All Specialties, United States (in millions)

	Flexible Sigmoidoscopy (CI)	Colonoscopy (CI)
Current Capacity	3.0 (±0.4)	15.5 (±2.3)
Potential Capacity	10.3 (±1.1)	24.4 (±2.6)
Unused Capacity	**7.3**	**8.9**

Assuming 50-week working year

4

SLIDE NOTES 4: These results are collapsed across all specialties. (We do have results by specialty; they will be included in our final study report.) The results are also scaled up from the weekly assessments to annual estimates assuming a 50-week working year.

Respondents indicated that in a year, about 3 million flexible sigmoidoscopies are being done, and 15.5 million colonoscopies are being done. We are currently referring to that as "current capacity"; though, it might better be called "current utilization."

Then the potential capacity is the number that the respondents said they could do with their current resources. That estimate gives 10.3 million sigmoidoscopies and 24.4 million colonoscopies.

Note, however, that the questions about potential capacity were asked independently for each modality. Therefore, we cannot be certain that the respondents could reach the estimated potentials for both sigmoidoscopy and colonoscopy simultaneously.

SLIDE 5

Number of People requiring CRC Screening and Follow-up Procedures (in millions)
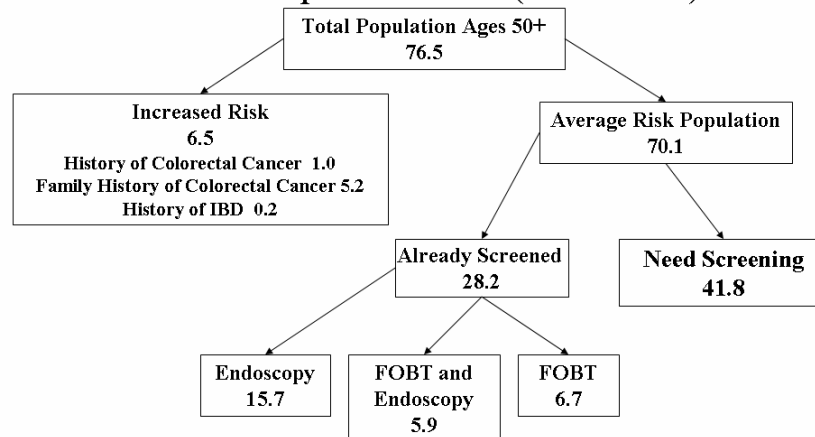
SLIDE 5 NOTES: We attempted to determine the size of the U.S. population that has not been screened and the number of tests needed under a series of scenarios to screen those people.

From the 76.5 million individuals in the U.S. age 50 or greater, we removed 6.5 million persons who would be deemed "high risk" from these categories. High-risk is based on family history of colorectal cancer or history of pre-disposing conditions. Those data come from SEER (NCI), National Health Interview Survey (CDC), and National Institute of Diabetes and Digestive and Kidney Diseases (NIH) databases.

The average-risk population is 70 million. We used NHIS data to determine which of those individuals would already have been screened, which eliminates 28.2 individuals. Thus, almost 42 million people ages 50 and older have not been screened according to current screening guidelines.

SLIDE 6

Distribution of Tests to 41.8 Million Unscreened: Hypothetical Screening Programs

- Base Case: current screening patterns
- Option 1: 100% FOBT, follow-up with colonoscopy
- Option 2:
  - 50% FOBT only
  - 25% FOBT plus sigmoidoscopy
  - 25% colonoscopy
- Option 3: 100% colonoscopy

6

SLIDE 6 NOTES: We looked at how to distribute the current endoscopy capacity among the 42 million unscreened individuals. Several options are available. First, we examined how capacity would compare with unmet screening need if the unscreened population were to follow current screening patterns. We assume that those 42 million people would be screened in the same proportions as is currently observed through the National Health Interview Survey (NHIS).

Option 1 assumes that all those people would get FOBT first, followed by a follow-up colonoscopy for those individuals with a positive FOBT.

Option 2 assumes that half of the individuals would get an FOBT, one-quarter would get FOBT plus flexible sigmoidoscopy, and one-quarter would get colonoscopy.

Option 3 assumes that all such individuals would be screened via colonoscopy.

SLIDE 7

Number of CRC Screening and Follow-up Tests

Required to Satisfy Unmet Need (in millions)

	FOBT	FSIG	Colonoscopy		
			Screening	Follow-up	Total
Base Case	18.5	16.1	15.2	1.2	16.3
Option 1	41.8	NA	NA	1.0	1.0
Option 2	31.3	10.5	10.5	1.2	11.6
Option 3	NA	NA	41.8	NA	41.8
Option 4	1.7	0.6	0.6	.06	0.6

Base Case: Current screening patterns
Option 1:   100% FOBT
Option 2:   Mixed tests
Option 3:   100% colonoscopy
Option 4:   Mixed tests; limited population

7

SLIDE 7 NOTES: Here are the results under each of the options. These results do not include surveillance colonoscopies that would be called for as part of an appropriate screening regimen. Our full study will examine how such guidelines would affect the estimates of need.

SLIDE 8

Capacity to Provide Lower Endoscopy: Comparison of
Unused Capacity to Unmet Need (in millions)

	Flexible Sigmoidoscopy				Colonoscopy			
	1 year	2 years	3 years	4 years	1 year	2 years	3 years	4 years
Base Case	**-8.8**	**-0.7**	1.9	3.3	**-7.7**	0.2	2.9	4.2
Option 1	NA	NA	NA	NA	7.6	7.6	7.6	7.6
Option 2	**-3.2**	2.1	3.8	4.7	**-3.1**	2.4	4.2	5.1
Option 3	NA	NA	NA	NA	**-33.2**	**-12.3**	**-5.3**	**-1.8**

Base Case: Current screening patterns

Option 1:   100% FOBT

Option 2:   Mixed tests   8

Option 3:   100% colonoscopy

SLIDE 8 NOTES: This table shows how the unused capacity (resulting from the SECAP) survey compares to the unmet need results if we were to distribute tests over 1, 2, 3, or 4 successive years. It would, of course, be difficult to do all tests in the first year.

As you can see, in the base case scenario, distributing the tests over a single year would leave a capacity deficit of almost 9 million sigmoidoscopies and almost 8 million colonoscopies. However, under option 1, we would have enough colonoscopy capacity to screen all individuals immediately.

If we were to try to offer colonoscopy to the entire unscreened population (option 3), it would take 5 years in all.

SLIDE 9

Conclusions

- Endoscopic capacity for general population immediately available for Option 1 (100% FOBT)

- Endoscopic capacity not immediately available for base case (current screening test patterns), Option 2 (mixed procedures), or Option 3 (100% colonoscopy)
  - Test would need to be distributed over 2 or more years

9

SLIDE 9 NOTES: Assuming that the current met need continues to be met, the endoscopic capacity for the general population is immediately available if we were to use an FOBT-only strategy. Any of the other options would require more than one year for completion.

REFERENCES

Seeff LC, Nadel MR, Klabunde CN, Thompson T, Shapiro JA, Vernon SW, Coates RJ. 2004. Patterns and Predictors of Colorectal Cancer Test Use in the Adult U.S. Population. *Cancer*. 100(10): 2093–2103.

Appendix O

Colorectal Cancer Surveillance
Testing After Polypectomy
Deborah Schrag, M.D., M.P.H.

SLIDE 1

Appendix O
Colorectal Cancer
Surveillance Testing After Polypectomy

Deborah Schrag MD MPH

Assistant Professor of Medicine and
Public Health Memorial Sloan-Kettering
Cancer Center

New York New York

1

SLIDE 1 NOTES: No notes.

SLIDE 2

Overview

- What is the evidence base supporting surveillance after diagnosis of an adenoma ("SAA) ?

- What do the guidelines recommend?

- What happens in real world practice?

- Possible sources of variation in CE Model outcomes

2

SLIDE 2 NOTES: No notes.

SLIDE 3

What is the Evidence Base
Supporting Strategies for SAA?

3

SLIDE 3 NOTES: No notes.

SLIDE 4

National Polyp Study
DESIGN

- Randomized trial of timing of surveillance colonoscopy
- ~1400 patients with resected adenomas
- Patient accrual 1980-1990 at 7 centers
- All polyps removed
- Central pathology review for all polyps
- Cecum reached; location and size of all polyps defined

4

SLIDE NOTES 4: Any discussion of surveillance would be remiss if it did not focus extensively on the National Polyp Study, which I will refer to as the NPS (Winawer et al., 1993).

The NPS was a randomized trial of the timing of surveillance colonoscopy. It included over 1400 patients who had resected adenomas. The seven participating centers were centers of excellence and expertise.

SLIDE 5

National Polyp Study
Surveillance Arms Following Randomization
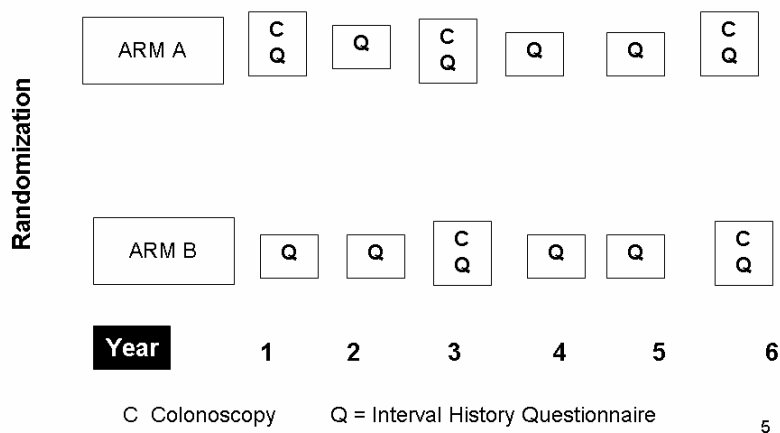
SLIDE 5 NOTES: The basic design involved a comparison of two arms. A patient randomized to Arm A was assigned to a surveillance colonoscopy at 1 year, at 3 years, and at 6 years following the colonoscopy in which the index adenoma was detected. In each of the years between the surveillance procedures, questionnaires were administered.

SLIDE 6

NPS Outcomes at Surveillance

- **Any adenoma**
- Advanced adenoma (NPS definition)
  - High grade dysplasia
  - Invasive cancer
  - Size > 1.0 cm
- Advanced adenoma (Current)
  - High grade dysplasia
  - Invasive cancer
  - ➢ Villous component
  - ➢ Size > 1.0 cm

6

SLIDE 6 NOTES: Outcomes of each surveillance procedures were defined as shown above. An advanced adenoma was defined in the study somewhat differently from that used in today's conventional current practice and in the guidelines issued by professional societies. Notice the subtle difference in the size criterion. There is some evidence that specifying greater than or equal to, as opposed to greater than, 10 mm can make a real non-trivial difference in the number of lesions subject to surveillance. In addition, a villous component has been recognized as an important risk indicator, but the data from the NPS did not define advanced adenomas using that criterion.

SLIDE 7

National Polyp Study
Findings at Follow-Up

	Arm B	Arm A		
	Three Year Only	One Year Only	One Year and Three Year	Three Year Given One Year
Adenoma	32.0%	27.5%	41.7%	21.6%
Advanced Adenoma	3.3%	2.6%	3.3%	0.9%

7

SLIDE 7 NOTES: Here are the results for the two arms. Arm B – the three-year surveillance protocol – had a 32 percent incidence of any adenoma and a 3.3 percent incidence of advanced adenoma. Arm B, on the other hand, had a higher cumulative incidence of adenomas (41.7 percent). However, the risk of advanced adenoma at three years is identical between the two arms (3.3 percent).

At the end of 6 years, the cumulative risk of any adenoma was 52 percent for Arm A and 48 percent for Arm B.

SLIDE 8
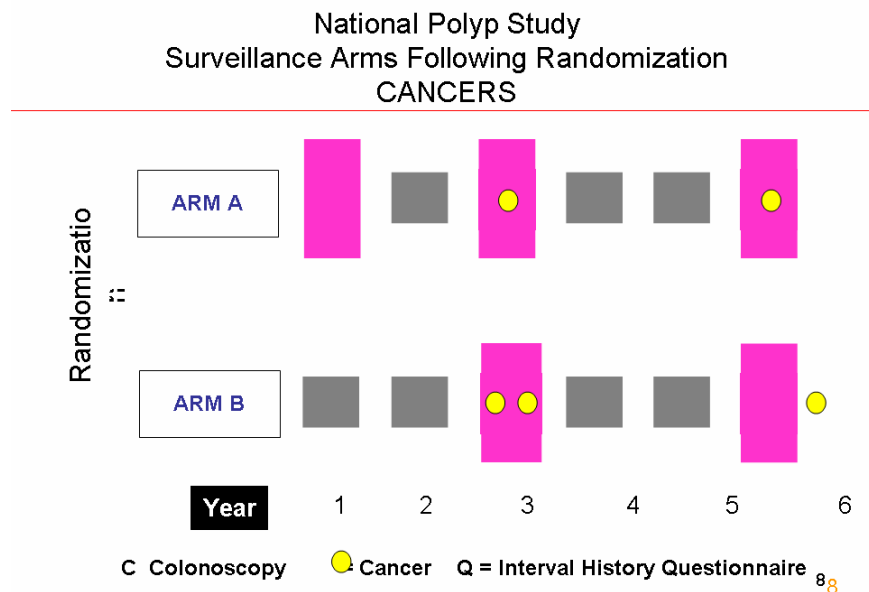
SLIDE 8 NOTES: This chart shows the point of detection of 5 cancers detected in the study population. The detection of cancers appears not to have been affected by the surveillance strategy.

SLIDE 9

**National Polyp Study
Cumulative Incidence of Adenomas with Advanced Pathology in
938 Patients by Randomization Arm**
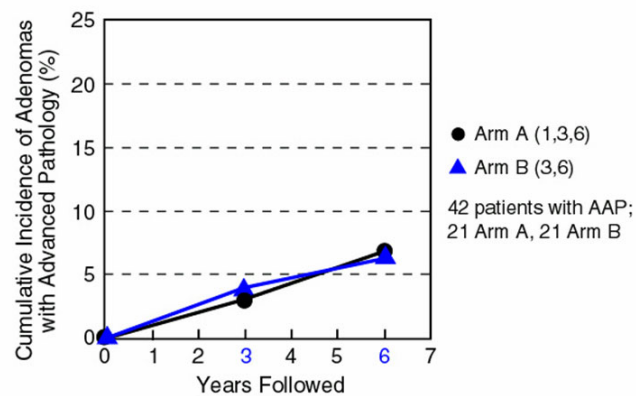
SLIDE 9 NOTES: The present and next 2 charts show findings of advanced adenomas (as defined by the NPS) as a function of surveillance strategy and various risk categories.

This chart shows that for advanced adenomas, the difference in incidence between the two arms is not different.

SLIDE 10

**National Polyp Study
Cumulative Incidence of Adenomas with Advanced Pathology
by Age at Diagnosis and Parental History of Colorectal Cancer**
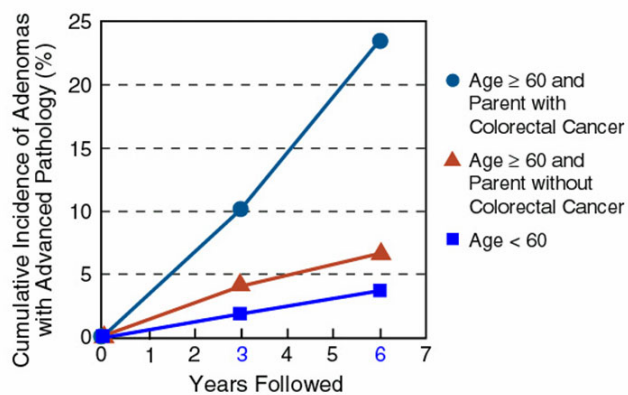
10   10

SLIDE 10 NOTES: Stratifying patients by age and status of a parent with colorectal cancer shows the increased probability of an advanced adenoma in the years following diagnosis. The rate for patients 60 years of age or older with a parent with colorectal cancer is almost 25 percent.

SLIDE 11

National Polyp Study
Cumulative Incidence of Adenomas with Advanced Pathology
by High and Low Risk Groups
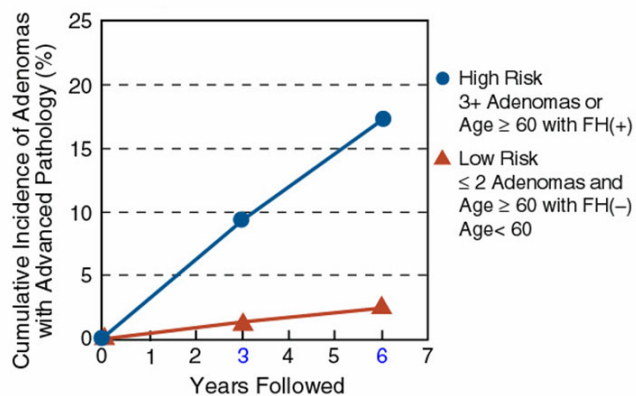
SLIDE 11 NOTES: If a person had 3 or more high-risk adenomas at the index event, the risk of an advanced adenoma by the sixth years of surveillance is over 15 percent.

SLIDE 12

National Polyp Study
Recommendations

- One year surveillance interval can be eliminated. (Winawer, et al., 1993)

- Surveillance intervals lengthened to 6 or more years for low risk adenoma patients (Zauber, et al., 1997)

12

SLIDE 12 NOTES: On the basis of these findings, the NPS researchers concluded that surveillance intervals lengthened to six or more years are reasonable for low-risk adenoma patients. (Winawer et al., 1993; Zauber and Winawer, 1997).

The data are currently being reanalyzed with the current definition of high-risk adenoma. It is unlikely, however, that the reanalysis will affect the conclusions.

SLIDE 13

Cancer Incidence Observed in Post Polypectomy Prevention Trials

Cohort	N	Person-yrs of observation	CRC/PYO x 10^{-3}
NPS	1418	8400	0.6%
Schatzkin	1905	5810	2.4%
Alberts	1303	3789	2.4%
Van stolik	930	3447	2.3%
Bertario	1063	8906	1.2%

13

13

SLIDE 13 NOTES: Is it possible that cancers were missed in the NPS? Some relatively small studies have examined back-to-back colonoscopies. They have shown that the miss rate is inversely related to the size of the tumor. For advanced adenomas, the miss rate is small, so it is unlikely that the clinicians in the NPS missed a high number of tumors.

SLIDE 14

Adenoma Miss Rates in Tandem Colonoscopies

Study	N	<6mm	6-9mm	AA
Hixon	90	16%	12%	0
Rex	183	27%	13%	6%

14 14

SLIDE 14 NOTE: Is it possible that cancers were missed in the NPS? Some relatively small studies have examined back-to-back colonoscopies. They have shown that the miss rate is inversely related to the size of the tumor. For advanced adenomas, the miss rate is small, so it is unlikely that the clinicians in the NPS missed a high number of tumors.

SLIDE 15

What Do Professional
Organizations and Societies
Recommend Regarding
Surveillance Following Removal
of an Adenoma?

15

SLIDE 15 NOTES: No notes.

SLIDE 16

Recommendations for SAA

- GI MultiSociety Task Force   Consortium 2003
  - American Society of Gastrointestinal Endoscopy
  - American Gastroenterological Association
  - American College of Physicians
  - Society of General Internal Medicine
  - American Academy of Family Practitioners
  - American Society of Colorectal Surgeons
  - Update of AHCPR (AHRQ) 1997 Panel
- American College of Gastroenterology ACG
- American Cancer Society   ACS

16 16

SLIDE 16 NOTES: There are currently three sets of recommendations for surveillance after adenomas. The multi-society task force, which issued guidelines in 2003, represents the consensus of a number of different provider groups.

SLIDE 17

SAA: Guideline Comparison

Clinical Scenario	Multi-Society 2003	ACG 2000	ACS 2001
1-2 SAs, no AA	5 yrs	5 yrs	3-6 yrs
1-2 SAs, no AA, +FH	-	3 yrs	-
After – FU	5 yrs	5 yrs	Average risk
		Maybe longer	(10 yrs)
Multiple SAs, no AA	3 yrs	3 yrs	3 yrs
After – FU	5 yrs	5 yrs	3 yrs, if - avg risk
AA	3 yrs	3 yrs	3 yrs
After - FU	5 yrs	5 yrs	3 yrs, if - avg risk

17

SLIDE 17 NOTES: The three groups are not very different from one another in their recommendations. After two small adenomas, or no adenoma with advanced pathology, most societies recommend surveillance after 5 years. The American Cancer Society gives a greater range – from 3 to 6 years. The groups do differ on what to do after the first negative surveillance examination. Two groups indicate that every-five-years is still warranted, whereas the American Cancer Society recommends that such individuals be considered average risk and returned to a screening pool (with a colonoscopy ever 10 years).

SLIDE 18

What are Population-Based Estimates of Current Surveillance Practice?

18

SLIDE 18 NOTES: No notes.

SLIDE 19

Population-Based Estimates
of SAA:

19

SLIDE 19 NOTES: I am not going to address this issue because Todd Anderson's presentation directly addresses it.

SLIDE 20

Adherence to Surveillance Guidelines

- Survey Literature suggests potential importance of various influences on physician behavior
  - Economic incentives
  - Referral networks/practice type
  - Last case bias
  - Medicolegal concerns
  - Guideline adherence as performance measure
  - MD specialty and experience
  - Pathologists degree of meticulousness

20 20

SLIDE 20 NOTES: No notes.

SLIDE 21

Risk is Heterogeneous

Results of Initial Colonoscopy	AA by 6 years
>=3 Advanced Adenoma, Age>60, FDR with CRC	20%
Age<60, 1 Adenoma	<1%

21

SLIDE 21 NOTES: It is important to remember, when considering adherence data, that risk is heterogeneous. We know from the National Polyp Study, for example, that 20 percent of individuals who are over age 60 and have a first-degree relative with CRC are likely to have an advanced adenoma in the 6 years following the index polypectomy, but only 1 percent under age 60 with a single adenoma and no first-degree relative with CRC will have one. It appears, from the surveys, that physicians actually take those risk differences into account when they decide on what to recommend to their individual patients.

SLIDE 22

Model SAA Parameters

Author	Size Contingent	Schedule	AA Definition
Ness	No	Q3 yrs	-
Kuntz	yes	Q3 yrs	current
Vijan	Yes	Q5 if AA, otherwise not	-
Loeve	Yes	Q3 if AA Q5 if not	current
Ladabaum	No	Q 5yrs until 80	-

22

SLIDE 22 NOTES: A brief look at the five models investigated in depth as part of this workshop shows that the surveillance patterns differ substantially. Some base surveillance strategies contingent on the size of the initial adenoma, some don't. Most use a 3-year surveillance schedule. And they differ as to whether a different surveillance regimen is enacted for advanced adenomas.

What none of the models appear to have is any tailoring of surveillance on the basis of age. But the data suggest that in practice, age is an important determinant of the frequency of surveillance.

REFERENCES

Winawer SJ, Zauber AG, May Nah Ho, O'Brien MJ, Gottlieb LS, Sternberg SS, Waye JD, Schapiro M, Bond JH, Panish JF, Ackroyd F, Shike M, Kurtz RC, Hornsby-Lewis L, Gerdes H, Stewart ET, Lightdale CJ, Edelman M, Fleisher M. 1993. Prevention of colorectal cancer by colonoscopic polypectomy. *N Engl J Med.* 329(27): 1977–1981.
Zauber AG, Winawer SJ. 1997. Initial management and follow-up surveillance of patients with colorectal adenomas. *Gastroenterol Clin North Am.* 26(1): 85–101.

Appendix P

Natural History of Colorectal Adenomas and Cancer
T. R. Levin, M.D.

SLIDE 1

Appendix P
Natural History of Colorectal
Adenomas and Cancer

T. R. Levin, MD
Kaiser Permanente,
Northern California

1

SLIDE 1 NOTES: No notes.

263

SLIDE 2

Topics

- The Polyp → Cancer Sequence
  - *De Novo* Cancers vs. From Pre-existing Adenoma
  - Distribution of Dwelling Times from Clinically Detectable Adenoma to Cancer
  - Incidence and Prevalence of Polyps in Average Risk Americans by age, location, type
  - Prognostic value of polyp detection for synchronous and metasynchronous polyps or cancers
- Stage specific dwell times for cancers

2

SLIDE 2 NOTES: No notes.

SLIDE 3

De Novo Cancer vs. Pre Existing Adenomas

- Clinically detectable vs. biologically pre-existing
  - Flat adenomas, Serrated adenomas, polypoid adenomas that are missed endoscopically
- Received Wisdom: "10%"
- Case-cohort Markov model (Taiwan)
  - 27-32% of cancers are De Novo
- Case control studies of SIG:
  - 60% reduction in CRC mortality for distal cancer
  - ?40% of distal cancers *De Novo* ?

3

SLIDE 3 NOTES: The concept of de novo cancer differs according to the perspective of the individual using the term. To an endoscopist it may mean one thing; to a molecular biologist another. For the purpose of an outcome study or a CEA model of colorectal cancer screening, it may not matter. Any adenoma that cannot be found on any of the existing technologies can be referred to as de novo. However, over time, colonoscopy and other techniques may improve our ability to detect lesions that cannot be detected today.

Before I re-reviewed the literature on this subject, I believed that the general consensus for the proportion of cancers that arise de novo -- 10 percent – was substantially correct. However, one case-cohort Markov model from Taiwan followed about 13,000 people who had colonoscopy and recorded the percent who developed colon cancer during followup (Chen et al., 2003). They estimated a very high rate of 27-32 percent, depending on assumptions such as the detectability of adenomas on colonoscopy. Although the population was Taiwanese, which could have a different natural history or different endoscopic techniques, I believe this is an upper bound and possibly an overestimate.

Kaiser's study found a 60 percent mortality reduction for distal cancer from people exposed to rigid sigmoidoscopy (Selby et al., 1992). Can that study be interpreted to suggest that 40 percent of distal cancers arise de novo? Of course that study reflected a series of examinations given to all comers, regardless of the quality of the examination, degree of penetration of the sigmoidoscope, etc. Therefore, many of the cancers do not meet the true definition of de novo. Rather, they reflect the existence of endoscopic failures.

SLIDE 4

National Polyp Study

Age at enrollment (yr)	72	68	72	60	76
Sex	F	F	F	F	M
No. of examinations in first 3 yr	2	1	1	1	2
Most advanced pathological finding at enrollment†	Villous adenoma	Tubular adenoma	Tubular adenoma	Tubular adenoma	Villous adenoma
Largest adenoma at enrollment (cm)	2.5	0.3	0.3	1.5	3.0
No. of adenomas at enrollment	3	1	1	3	4
Grade of dysplasia at enrollment†	High	Low	Low	Low	Low
Year of follow-up when malignant polyp was detected	6	7	3	3	3
Size of malignant polyp (cm)	0.6	0.8	2.5	1.5	1.5
Location of malignant polyp	Sigmoid colon	Hepatic flexure	Transverse colon	Cecum	Cecum
Surgery	No	Yes	Yes	Yes	Yes
Stage					
TNM	Indeterminate	T1N0M0	T1N0M0	T1N0M0	T2N0M0
Dukes	Indeterminate	A	A	A	B
Current status	Alive	Alive	Alive	Alive	Alive

5 of 1418 = **0.35%**, 7 yrs. f/up, = **0.05%/person-yr** [4]

SLIDE NOTES 4: The National Polyp Study found five cancers as part of surveillance colonoscopies (Winawer et al., 1993). The first one listed was quite small and was detected only by an expert colonoscopist in the sigmoid colon. The others were all beyond the reach of sigmoidoscopy and might be classified as left-sided cancers.

When these findings are compared with the Kaiser sigmoidoscopy findings (previous slide) they raise the question whether the only cancers that can be prevented are ones in the distal colon.

SLIDE 5

PLCO

Table 4. Neoplastic Findings 3 Years After Negative Sigmoidoscopy (N = 9317)*

	Distal Colon	Proximal Colon†	Cum
Nonadvanced adenoma	214	124	
Advanced adenoma	72	39	
Cancer	6	1	
Total	292 (3.1)	164 (1.8)	

*Data are presented as No. or No. (%).

6 of 9317 = **0.06%** for 3 years, or **0.02%** per person-year

5

SLIDE 5 NOTES: Here is a table from the PLCO study (Schoen et al., 2003). This paper suggests an incidence of de-novo distal cancer of 0.02 percent per year, at least in patients undergoing sigmoidoscopy. This result is not much different from the findings of the National Polyp Study.

SLIDE 6

Adenoma vs. *De Novo*

- "Natural History" follow-up of KP CoCaP
- Incidence based on clinical Dx of Cancer, not follow-up colonoscopy
  - 72,483 avg. risk members with a negative screening sigmoid (1994-1996; 4 yrs. f/up).
  - 30 cases of distal cancer; 302,424 person-years of follow-up (**0.01% per person-year**)

6

SLIDE 6 NOTES: Paul Doria-Rose has been working on a study at Kaiser Permanente that looks at early cohorts from our sigmoidoscopy screening program (Doria-Rose et al., 2004).

Focusing on the distal cancers only, about 30 cases occurred over 302,424 person years of follow-up, or approximately 0.01 percent per person year. That rate is somewhat lower than was reported in the PLCO trial, probably because the individuals in this study developed diagnosed cancer, whereas in PLCO some cancers were found through re-screening.
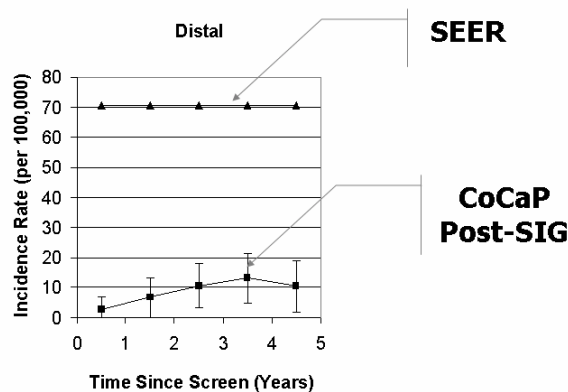
SLIDE 7

SLIDE 7 NOTES: We compared the experience in the Kaiser sigmoidoscopy program with the incidence of cancer reported in the SEER registry for our region. The comparison implies about an 85 percent reduction in incidence of CRC during those years of follow up (Doria-Rose et al., 2004). Would that imply that roughly 15 percent of the cancers are de novo?

SLIDE 8

Telemark: FS to Select for Colonoscopy (N = 799)
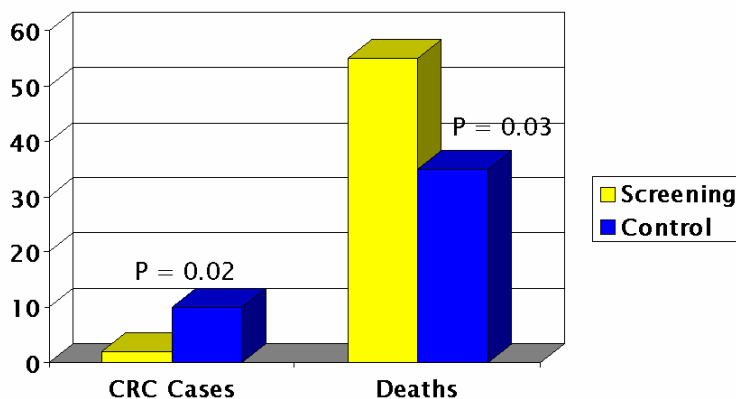
80% reduction in CRC incidence: ?20% "De Novo?"   8

SLIDE 8 NOTES: The Telemark randomized trial compared sigmoidoscopy vs. no screening, and then polypectomy at colonoscopy and colonoscopy surveillance (Thiis-Evensen et al., 1999). Two interesting results from that study may be relevant. One is the 80 percent reduction in colorectal cancer incidence over the follow-up period. Does that mean that we cannot prevent 20 percent of cancers because they are effectively de novo at present? Or does it mean that sigmoidoscopy cannot detect many cancers?

Another interesting finding of this study is that there were actually more deaths in the screening group than there were in the control group. Any mortality reduction that was achieved with colorectal cancer was completely swamped by excess cardiovascular death. One theory is that somehow screening is actually contributing to cardiovascular disease and death. But this is just hypothesis. Given the theories about the relationship between infectious disease and cardiovascular disease, perhaps we are liberating some microbes from the colon at the time we insert a scope. Such a hypothesis would be a worst-case scenario, and at present there is simply no evidence to explain these findings.

Whether or not colorectal cancer screening increases or simply does not decrease all-cause mortality, we must ask whether by screening we simply save people from colorectal cancer to die of other things.

SLIDE 9

Adenoma vs. *De Novo*

- Still to come:
  - Colonoscopy screening follow-up studies
    - VA Coop #380
  - Consensus estimates
    - R. Ness, Vanderbilt
- Proportion probably varies with age, gender, location, race/ethnicity

9

SLIDE 9 NOTES: Other sources of data on de novo rates are expected soon, from David Lieberman's cooperative study of follow-up from colonoscopy screening, and the consensus estimates obtained by Reid Ness.

Finally, the true proportion of de novo cancers probably depends on the age and gender of the person. They may be more common in women, especially as we are learning that proximal cancers appear to be more common in women and de novo cancers are more likely to be proximal cancers. African Americans may also have a higher rate of de novo cancers than do Caucasians.

SLIDE 10

Adenoma to Cancer Dwell Times

- 1990's: "10 years."
  - Direct observation of biopsied polyps left in place (10-15 years)
  - Cancer incidence after excision of polyps
    - (14 years)
  - Case control evidence of protection > 10 years
    - Now up to 14 years (P. Newcomb)

10

SLIDE 10 NOTES: The next issue is what is the best estimate of the length of time it takes for adenomas to transition into cancer. One problem in attempting to estimate this is the fact that there may be two kinds of adenomas – those that are never going to progress and those that will progress. We are just beginning to scratch the surface in understanding what influences the progression rate.

The "10 year" assumption comes from studies of biopsied polyps that were left in place (Morson, 1984). Those very few studies found that it took about 10 to 15 years before cancer developed.In a study of cancers in patients with polyps removed, Atkin found an average 14 year lapse (Atkin et al., 1992). Finally, case control studies, including a recent one by Polly Newcomb, show that the protection persists up to 14 years (Selby et al., 1992; Newcomb et al., 2003). So, the evidence from different studies is consistent and points to a protective period of about 14 years.

SLIDE 11

Distribution of Dwell Times

– 2.5 adenomas/1000 per year progress to cancer
– Barium Enema Study (Stryker) large polyps→ cancer
  • 2.5%, at 5 years, 8% at 10 years, 24% at 20 years
– Is there anything new since 1987?

11

SLIDE 11 NOTES: I found no recent information on the distribution of dwelling times. The Stryker study is the most frequently cited (Stryker et al., 1987). That Mayo Clinic report suggests that about 75 percent of large polyps found on barium enema and left unresected for various reasons never progressed to cancer in the 20-year period of observation .

A study of small adenomas showed that about 2.5 adenomas per 1,000 per year progressed to cancer (Eide, 1986).

SLIDE 12

Prevalence of Polyps:
Age, Location, Type

- Adenomas in 25% of people by age 50; large adenomas 4.6% at 54, 15.6% at 75 years (based on autopsy data)
- VA Cooperative Study
- Lilly
- KP

12

SLIDE 12 NOTES: The guidelines promulgated by the GI Consortium led by Dr. Winawer concluded that adenomas of any size were present in 25 percent of people by age 50, large adenomas in 4.6 percent by age 54, and large adenomas in 15.6 percent of patients at age 75 (Winawer et al., 1997). These conclusions were based on a reading of the autopsy literature, which has been referred to by other presenters at this Workshop.

Probably the three largest studies that have looked at this in living patients (with endoscopy) have been the VA Cooperative Study, the Lilly Colonoscopy Study by Tom Imperiale, and our study at Kaiser (Levin et al., 1999; Imperiale et al., 2000; Lieberman et al., 2000).

SLIDE 13

Advanced Proximal Neoplasia by Distal Findings at Sigmoidoscopy

Worst Distal Finding	CoCaP 1999	VA Coop 2000	Lilly 2000
No adenomas	5.3%	2.7%	1.1-2%
Tubular Adenoma < 1cm diameter	5.5%	6.4%	4%
Tubular Adenoma ≥ 1cm diameter	5.7%	8.6%	4%
Villous Adenoma, any size	12.2%	12.5%	4%

13

SLIDE 13 NOTES: Regarding the question of the frequency with which advanced proximal neoplasia – both adenomas and cancer – are accompanied by a synchronous adenoma in the distal colon, there are a number of studies that suggest the rate is about 4 to 10 percent of people with a small distal tubular adenoma, and 1 to 5 percent in people without an adenoma in the distal colon.

The Kaiser CoCaP study found 5.3 percent of individuals with no adenomas on sigmoidoscopy had at least one advanced proximal adenoma, but most of these were not cancer (Levin et al., 1999). Our study was prospective, so we were not dependent on retrospective review of electronic and paper medical records, as some studies were.

SLIDE 14

KP-CoCaP Data (CART Analysis)
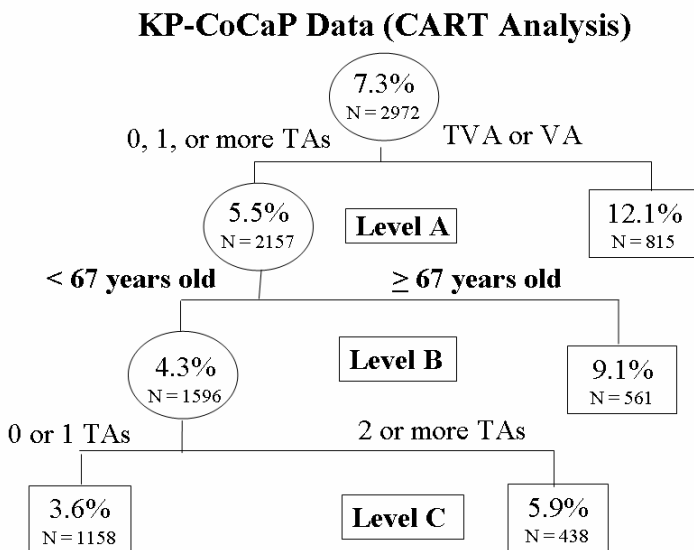
%'s are the prevalence of APN in each subgroup

14

SLIDE 14 NOTE: As part of the Kaiser study we analyzed (through classification and regression analysis) the risk factors associated with advanced proximal neoplasia (adenomas and cancer) by type in close to 3,000 people who had colonoscopy (Levin et al., 1999). The most important risk factor was the presence of a villous architecture in the polyp; 12.1 percent of those individuals were found to have advanced proximal neoplasia. The next most important factor was age. Notice, also, that the number of tubular adenomas is not as important as the other two. Indeed, age alone is a stronger predictor than the presence of a distal tubular adenoma. Distal adenoma size did not enter into the CART analysis at any level.

These findings suggest the potential for mixed strategies – using different screening tests at different ages – to optimize a screening program. For example, perhaps it would be better to screen younger people with sigmoidoscopy and switch to a total colon examination after 60 or 65 years of age. This is where modeling could be especially useful, because it would take many years to answer the question via a clinical trial of such strategies.

SLIDE 15

FS vs. Colonoscopy: Sensitivity for Advanced Proximal Neoplasia

	Using Lower Cutoff (sigm/desc colon)	Using Higher Cutoff (splenic flexure)
Lieberman et al, NEJM 2000	68.1%	79.9%
Imperiale et al, NEJM 2000		72%

15

SLIDE 15 NOTES: This chart shows recent findings on the sensitivity of sigmoidoscopy for advanced proximal polyps and cancers (Imperiale et al., 2000; Lieberman et al., 2000). The two studies indicate that about 60-70 percent of clinically important lesions in the proximal colon can be found through following patients with lesions in the distal colon. That finding may vary with the reach of the sigmoidoscope into the colon, and it may also vary with age of the patient. Sensitivity could be lower in older patients, because of the higher prevalence of proximal lesions in that group.

SLIDE 16

Isolated Advanced Proximal Neoplasia
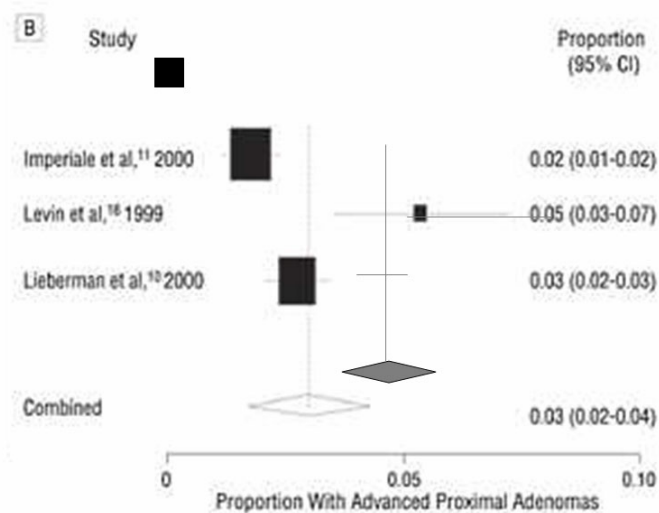
SLIDE 16 NOTES: Lewis did a meta-analysis of studies that examined the relationship between distal and proximal advanced adenomas and cancers. From the three studies, it appears that roughly 3 percent of all screened individuals have an isolated proximal adenoma that would not be detected by sigmoidoscopy. That implies that colonoscopy provides an incremental benefit over sigmoidoscopy to only 3 or so percent of the screened population (Lewis et al., 2003).

SLIDE 17

Prognostic Value of Polyps

- UK 14 year follow-up study:
  - Distal advanced adenomas (large or villous) increase risk of subsequent cancer 3 fold;
  - multiple advanced adenomas, increased 6-fold;
  - single or multiple small polyps predicted no increased risk
- Dietary intervention studies
  - Index adenomas to predict subsequent advanced neoplasia (cancers are rare)

17

SLIDE 17 NOTES: What is the prognostic value of finding adenomas? In Atkin's study of follow-up after sigmoidoscopy, the presence of a distal advanced adenoma increases the risk of subsequent cancer 3-fold; and multiple advanced adenomas increased it by 6-fold (Atkin et al., 1992). However, finding the presence of a single polyp or multiple small polyps did not carry an increased risk. (The relatively young age of the people in that study, however, may have an effect on these findings.)

The dietary intervention studies, which involve periodic surveillance with colonoscopy, also show that individuals with one or two small adenomas rarely get subsequent cancers.

SLIDE 18

Stage Specific Dwell Times for Cancer

– Clinically latent cancer is a real phenomenon
– Malignancy → clinical appearance 4.8 years
  • R. Koretz, Ann Int Med (malignant polyps)
– Disease Progression is heterogeneous

18

SLIDE 18 NOTES: Is there any new information that can inform modelers about the stage-specific dwelling times for colon cancer? Koretz's 1993 analysis showed that the time from the development of malignancy until clinical detection is close to five years. That analysis is based on a comparison of autopsy data and clinical reports of prevalence of cancers at different stages (Koretz, 1993).

SLIDE 19

Predicting Progression

- Stage (including micrometastases, CEA level)
- Histology: (serrated neoplasia, differentiation)
- Race/ethnicity, gender
- Molecular characteristics
  - MSI/TGF beta; 18q loss: *DCC, BRCA1, SMAD4*
  - Epidemiology of markers is not well known
  - Interact with stage, location, histology, race/ethnicity, gender
- Method of detection (occult blood, molecular marker, endoscopy) may interact with genotype, phenotype and prognosis.

19

SLIDE 19 NOTES: This slide includes a list of tumor and patient characteristics that appear to be correlated with progression of CRC. Note, however, that this list is by no means exhaustive. Other prognostic factors exist. The list shown here is from a single search of the medical literature (via Medline) conducted on 1-16-2004.

The molecular characteristics of tumors that affect both progression and response to therapy are not well understood. For example, we do not yet know what proportion of different cancers carry the various tumor markers that have been identified so far. These markers probably interact with stage at detection, location, histology, race, gender, and other more obvious and easily measured things. For example, cancers in African Americans may be more likely to have a specific molecular characteristic. However, at present we do not have enough information on how the molecular characteristics map into more easily measured characteristics.

I would propose that the method of detection of cancers and polyps may also interact with the genotype or phenotype or prognosis. It may be that cancers detected with stool marker tests that search for a certain molecular panel of fecal DNA may be different in their molecular distribution from those detected by more traditional fecal occult blood tests. And, those differences may mean that the progression of cancers found on such tests differs from those found with fecal occult blood.

SLIDE 20

Summary

- Most advanced adenomas will never develop clinically apparent cancer
  – Those that do, will take a long time
- Some cancers will arise "De Novo"
  – Biologic phenomenon? Endoscopy failure?
- Increasing age predicts proximal disease, interval cancers.
  – Molecular correlate: DNA methylation?
  – No reliable molecular correlate yet identified for villous histology

20

SLIDE 20 NOTES: Similarly, the cancers that can be seen with colonoscopy alone may be different in terms of progression from those that can be seen with both colonoscopy and virtual colonoscopy. Perhaps tumors located "behind a fold" (and therefore more difficult for virtual colonoscopy) may have a different biologic behavior. We simply do not know whether such is the case at present (Shibata et al., 1996; Carethers et al., 1998; Liefers et al., 1998; Watanabe et al., 2001; Marcella and Miller, 2001; Makinen et al., 2001; Liang et al., 2002; Garcia et al., 2003; McArdle et al., 2003; Rabeneck et al., 2003; Jemal et al., 2003; Kama et al., 2003).

We do know that most advanced adenomas will never develop into clinically apparent cancer. Those that do will take a long time. Some cancers will arise de novo, but we cannot know at present whether they are inherently unique biologically or whether they grew from adenomas that were missed on colonoscopy. This distinction may not matter unless there is one day a major advance in colonoscopic technology.

Increasing age does predict proximal disease, which probably leads to more cancers developing in the interval between screens. In the National Polyp Study, all individuals with interval cancers were over 60. There may be a molecular correlation with this phenomenon: DNA methylation is one cancer pathway that seems to accumulate as people age.

Finally, villous histology is such a strong predictor of cancer risk that there may be a molecular marker that correlates with it. However, to my knowledge no such marker has yet been identified.

REFERENCES

Atkin WS, Morson BC, Cuzick J. 1992. Long-term risk of colorectal cancer after excision of rectosigmoid adenomas. *N Engl J Med*. 326(10): 658-62.

Carethers JM, Hawn MT, Greenson JK, Hitchcock CL, Boland CR. 1998. Prognostic significance of allelic lost at chromosome 18q21 for stage II colorectal cancer. *Gastroenterology*. 114(6): 1188-95.

Chen CD, Yen MF, Wang WM, Wong JM, Chen TH. 2003. A case-cohort study for the disease natural history of adenoma-carcinoma and de novo carcinoma and surveillance of colon and rectum after polypectomy: implication for efficacy of colonoscopy. *Br J Cancer*. 88(12): 1866-73.

Doria-Rose VP, Levin TR, Selby JV, Newcomb PA, Richert-Boe KE, Weiss NS. 2004. The incidence of colorectal cancer following a negative screening sigmoidoscopy: implications for screening interval. *Gastroenterology*. 127(3): 714-22.

Eide TJ. 1986. Risk of colorectal cancer in adenoma-bearing individuals within a defined population. *Int J Cancer*. 38(2): 173-6.

Garcia JM, Rodriguez R, Dominguez G, Silva JM, Provencio M, Silva J, Colmenarejo A, Millan I, Munoz C, Salas C, Coca S, Espana P, Bonilla F. 2003. Prognostic significance of the allelic loss of the BRCA1 gene in colorectal cancer. *Gut*. 52(12): 1756-63.

Imperiale TF, Wagner DR, Lin CY, Larkin GN, Rogge JD, Ransohoff DF. 2000. Risk of advanced proximal neoplasms in asymptomatic adults according to the distal colorectal findings. *N Engl J Med*. 343(3): 169-74.

Jemal A, Murray T, Samuels A, Ghafoor A, Ward E, Thun MJ. 2003. Cancer statistics, 2003. *CA Cancer J Clin*. 53(1): 5-26.

Kama NA, Kologlu M, Reis E, Doganay M, Atli M, Dolapci M. 2003. A prognostic score for colorectal cancer. *Hepatogastroenterology*. 50(53): 1356-61.

Koretz RL. 1993. Malignant polyps: are they sheep in wolves' clothing? *Ann Intern Med*. 118(1): 63-8.

Levin TR, Palitz A, Grossman S, Conell C, Finkler L, Ackerson L, Rumore G, Selby JV. 1999. Predicting advanced proximal colonic neoplasia with screening sigmoidoscopy. *JAMA*. 281(17): 1611-7.

Lewis JD, Ng K, Hung KE, Bilker WB, Berlin JA, Brensinger C, Rustgi AK. 2003. Detection of proximal adenomatous polyps with screening sigmoidoscopy: a systematic review and meta-analysis of screening colonoscopy. *Arch Intern Med*. 163(4): 413-20.

Liang JT, Huang KC, Lai HS, Lee PH, Cheng YM, Hsu HC, Cheng AL, Hsu CH, Yeh KH, Wang SM, Tang C, Chang KJ. 2002. High-frequency microsatellite instability predicts better chemosensitivity to high-dose 5-fluorouracil plus leucovorin chemotherapy for stage IV sporadic colorectal cancer after palliative bowel resection. *Int J Cancer*. 101(6): 519-25.

Lieberman DA, Weiss DG, Bond JH, Ahnen DJ, Garewal H, Chejfec G. 2000. Use of colonoscopy to screen asymptomatic adults for colorectal cancer. Veterans Affairs Cooperative Study Group 380. *N Engl J Med*. 343(3): 162-8.

Liefers GJ, Cleton-Jansen AM, van de Velde CJ, Hermans J, van Krieken JH, Cornelisse CJ, Tollenaar RA. 1998. Micrometastases and survival in stage II colorectal cancer. *N Engl J Med*. 339(4): 223-8.

Makinen MJ, George SM, Jernvall P, Makela J, Vihko P, Karttunen TJ. 2001. Colorectal carcinoma associated with serrated adenoma--prevalence, histological features, and prognosis. *J Pathol*. 193(3): 286-94.

Marcella S, Miller JE. 2001. Racial differences in colorectal cancer mortality. The importance of stage and socioeconomic status . *J Clin Epidemiol*. 54(4): 359-66.

McArdle CS, McMillan DC, Hole DJ. 2003. Male gender adversely affects survival following surgery for colorectal cancer. *Br J Surg*. 90(6): 711-5.

Morson BC. 1984. The evolution of colorectal carcinoma. *Clin Radiol*. 35(6): 425-31.

Newcomb PA, Storer BE, Morimoto LM, Templeton A, Potter JD. 2003. Long-term efficacy of sigmoidoscopy in the reduction of colorectal cancer incidence. *J Natl Cancer Inst*. 95(8): 622-5.

Rabeneck L, Souchek J, El-Serag HB. 2003. Survival of colorectal cancer patients hospitalized in the Veterans Affairs Health Care System. *Am J Gastroenterol*. 98(5): 1186-92.

Schoen RE, Pinsky PF, Weissfeld JL, Bresalier RS, Church T, Prorok P, Gohagan JK. 2003 . Results of repeat sigmoidoscopy 3 years after a negative examination. *JAMA*. 290(1): 41-8.

Selby JV, Friedman GD, Quesenberry CP Jr, Weiss NS. 1992. A case-control study of screening sigmoidoscopy and mortality from colorectal cancer. *N Engl J Med*. 326(10): 653-7.

Shibata D, Reale MA, Lavin P, Silverman M, Fearon ER, Steele G Jr, Jessup JM, Loda M, Summerhayes IC. 1996. The DCC protein and prognosis in colorectal cancer. *N Engl J Med*. 335(23): 1727-32.

Stryker SJ, Wolff BG, Culp CE, Libbe SD, Ilstrup DM, MacCarty RL. 1987. Natural history of untreated colonic polyps. *Gastroenterology*. 93(5): 1009-13.

Thiis-Evensen E, Hoff GS, Sauar J, Langmark F, Majak BM, Vatn MH. 1999. Population-based surveillance by colonoscopy: effect on the incidence of colorectal cancer. Telemark Polyp Study I. *Scand J Gastroenterol*. 34(4): 414-20.

Watanabe T, Wu TT, Catalano PJ, Ueki T, Satriano R, Haller DG, Benson AB 3rd, Hamilton SR. 2001. Molecular predictors of survival after adjuvant chemotherapy for colon cancer. *N Engl J Med*. 344(16): 1196-206.

Winawer SJ, Fletcher RH, Miller L, Godlee F, Stolar MH, Mulrow CD, Woolf SH, Glick SN, Ganiats TG, Bond JH, Rosen L, Zapka JG, Olsen SJ, Giardiello FM, Sisk JE, Van Antwerp R, Brown-Davis C, Marciniak DA, Mayer RJ. 1997. Colorectal cancer screening: clinical guidelines and rationale. *Gastroenterology*. 112(2): 594-642.

Winawer SJ, Zauber AG, Ho MN, O'Brien MJ, Gottlieb LS, Sternberg SS, Waye JD, Schapiro M, Bond JH, Panish JF, et al. 1993. Prevention of colorectal cancer by colonoscopic polypectomy. The National Polyp Study Workgroup. *N Engl J Med*. 329(27): 1977-81.

Appendix Q

CISNET—
Cancer Intervention and Surveillance
Eric J. (Rocky) Feuer, Ph.D.

SLIDE 1

**Appendix Q
CISNET:
Cancer Intervention and Surveillance
Modeling Network**

Eric J. (Rocky) Feuer, Ph.D.

Chief, Statistical Research and Applications Branch, Division of
Cancer Control and Population Sciences, National Cancer
Institute

1

SLIDE 1 NOTES: No notes.

285

SLIDE 2

Overview

- **NCI Sponsored Consortium Focused on "Why" and "What If" Questions in Cancer Trends:**
  - ➤ Statistical Modeling of the Impact of Cancer Control Interventions (Screening, Treatment, Prevention) on Current and Future Trends
  - ➤ Optimal Cancer Control Planning
- **Two Rounds of Funding (4 year grants) – U01 Cooperative Agreement**
  - ➤ Sept. 2000 - 9 Funded Grants- Breast (7), Prostate (1) and Colorectal Cancer (1)
  - ➤ Aug. 2002: 8 Funded Grants - Prostate (1), Colorectal (2), and Lung Cancer (5)
  - ➤ RFA to be reissued shortly

2   NATIONAL CANCER INSTITUTE

SLIDE 2 NOTES: The CISNET initiative has some features that are closely aligned to the purposes of this workshop but many features in addition. It is an NCI-sponsored consortium focused on the "Why?" and "What if?" questions in cancer trends. So, it goes beyond the issues of cost-effectiveness modeling to the modeling of population trends.

CISNET is intended to further statistical modeling of the impact of cancer, taking into account the impact of the full range of cancer control interventions–prevention, screening and treatment – on current and future trends.

We have had two rounds of funding so far. These have been four-year grants by the National Cancer Institute (NCI).They are U01 cooperative agreements, which means that the staff at NCI work in cooperation with the grantees. In 2000, we began with three cancer sites – breast, prostate and colorectal cancer–and funded 9 grants. Most of those grants were in the area of breast cancer, reflecting the state of modeling at the time. Eight more grants were funded in 2002, and lung cancer was added to the list of sites

SLIDE 3
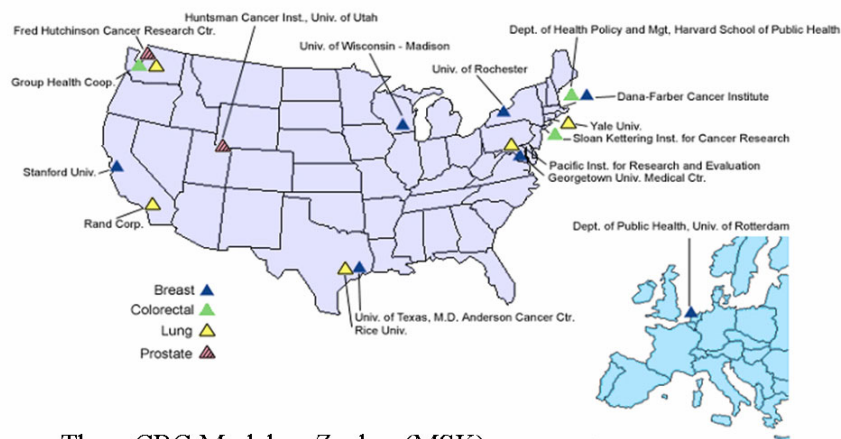
SLIDE 3 NOTES: The modelers in CISNET represent a well respected and diverse group of cancer modelers.

The MISCAN group from the Netherlands was one of the first groups in the world to do modeling of this type. Besides having a breast grant of their own, their colorectal model is being used by the funded group at Memorial Sloan Kettering, and the lung model being developed at Rand is a transplanted MISCAN person.

SLIDE 4

CISNET Goals

- **Develop and enhance multi-cohort based population models (including new methodology)**
  - ➢ Real cohorts representing the necessary birth cohorts to construct cross- sectional rates for specified years and age groups
  - ➢ Ideally can address the full range of interventions

- **Establish infrastructure to facilitate communication and understanding among modelers**
  - ➢ Model Profiler
  - ➢ Comparative Modeling Projects (Base Cases)

- **Gain access to data sets that might not otherwise be available**

- **Develop liaisons and provide assistance to outside groups to address questions amenable to modeling**
  - ➢ AHRQ/CMS – CE of immunochemical FOBT
  - ➢ CDC/NCI – Healthy People 2010 mid-course correction studies

4   NATIONAL CANCER INSTITUTE

SLIDE NOTES 4: No notes.

SLIDE 5

Types of Questions Population Models Can Address

- **Responsive to Challenges Due to Increasing Pace of Technology.**
  - ➢ Provide short term answers while Randomized Controlled Trials (RCT's) are still in progress (e.g. PSA testing for prostate cancer)

- **Address emerging questions while they are still being debated in the scientific/policy forum.**
  - ➢ Impact of smokeless tobacco products.
  - ➢ Virtual colonoscopy – esp. policies/practices in polyp-size threshold for sending patients on for conventional colonoscopy

- **Translate RCT evidence to the population setting.**
  - ➢ Population impact of adjuvant therapy and mammography for breast cancer.

- **Provide estimates of quantities that will never be derived from RCT's.**
  - ➢ Impact of smoking cessation in the US on future lung cancer trends.

5 NATIONAL CANCER INSTITUTE

SLIDE 5 NOTES: No notes.

SLIDE 6

Making Results of Modeling Efforts More Transparent

- **Modeling has been marred by:**

  ➢ Difficulty of understanding and comparing different model assumptions and structure

  ➢ Lack of comparability of inputs, outputs, basic definitions

- **Addressed with Model Profiler and Base Cases**

6   NATIONAL CANCER INSTITUTE

SLIDE 6 NOTES: This workshop grew out of one of the same issues that CISNET has had to address: the lack of comparability of inputs, outputs, structures and definitions among different models of the same cancer site.

We have tried to address this problem in CISNET with tools called "Model Profiler," and "Base Cases." I will discuss both of these efforts.
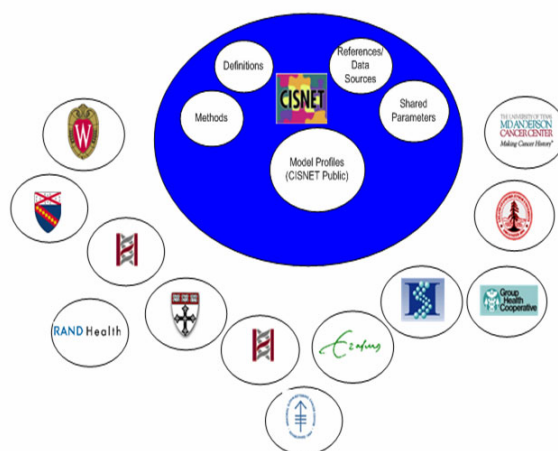
SLIDE 7

SLIDE 7 NOTES: Model Profiler is software that allows our collaborating centers to access a central CISNET interactive Web site that contains documentation on their own and other models in a consistent fashion.

SLIDE 8
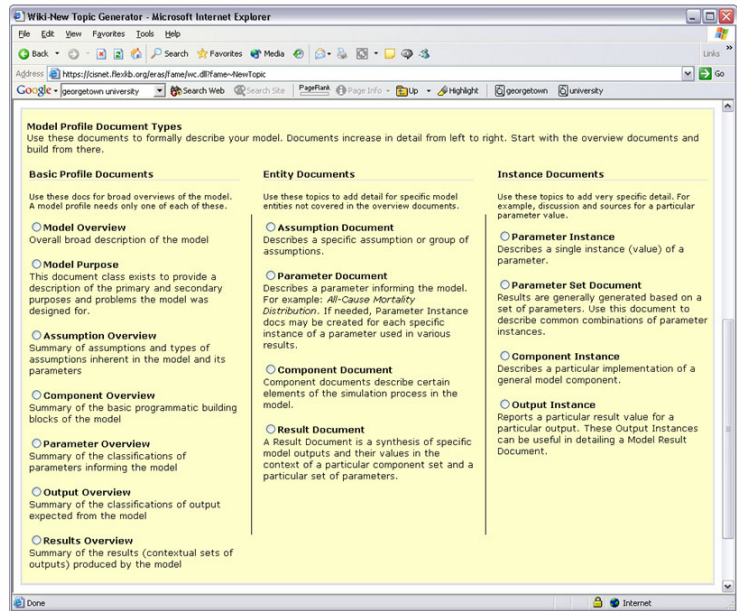
SLIDE 8 NOTES: Here is the structured format for entry of documentation into Model Profiler. There are a number of basic profile documents, a model overview, a model purpose, and an assumption overview. Ideally, all collaborators would enter information into each of these documents.

It is possible to go deeper and enter documents that address specific assumptions or parameters. Although we are encouraging groups to use this profiler, they are free to go into as much depth as they want to.
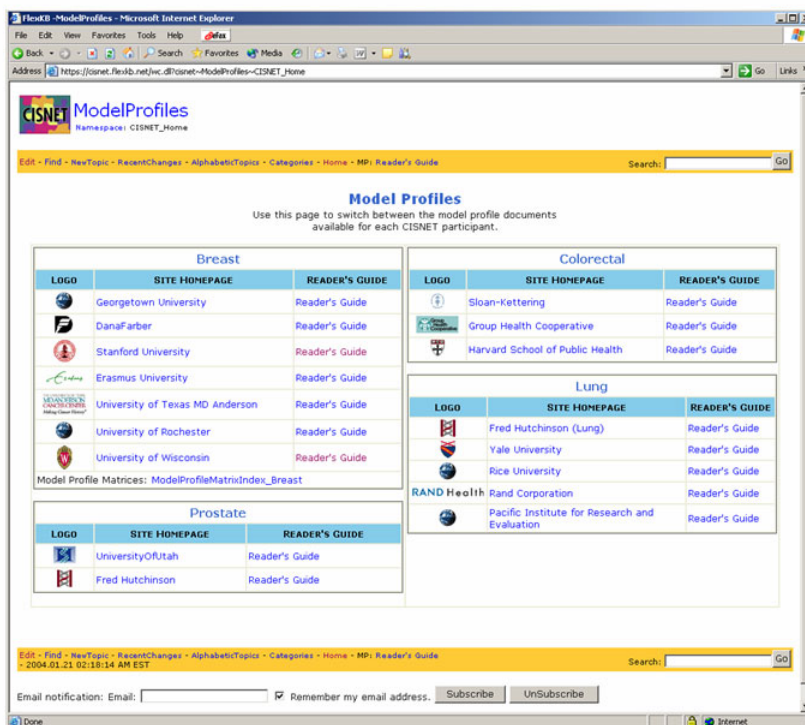
SLIDE 9

SLIDE 9 NOTES: When each group's data is posted on the central CISNET Web site, the table of contents page looks like this.

SLIDE 10

SLIDE 10 NOTES: Because each modeler answers the same set of questions, it is possible to compare the models relatively quickly. Here is an example of the Component Overview documents for a number of models.

SLIDE 11

SLIDE 11 NOTES: Here is an example of an analysis on breast cancer survival, where the rows are specific models and the columns represent certain characteristics of that model. This comparative analysis is done off-line by an analyst, but it can be posted in the Model Profiler.

SLIDE 12

SLIDE 12 NOTES: The final step, still to be taken, is to make the Model Profiler information available in a public site. The public site, cisnet.cancer.gov will be able to accommodate publications and documentation.

SLIDE 13

CISNET Base Cases

- Groups jointly decide to address a common question

- Common population-based inputs are used

- Modelers maintain the unique "deeper" aspects of their models (e.g. assumptions and formulation of the natural history of disease)

- Common computer runs and content/format of outputs are specified

- Provides a chance to reach common consensus on important questions, and to better understand differences between models

13   NATIONAL CANCER INSTITUTE

SLIDE 13 NOTES: The CISNET Base Cases initiative is similar to what the modelers at this workshop have been attempting to do. The group jointly decides to address a common question. We have common population-based inputs, and the modelers maintain what I would call the deeper aspects of their models, e.g., assumption and formulation of natural histories.

    We produce a set of common computer runs, and the format of the outputs is specified. That provides a chance to reach a consensus on important questions and to better understand differences among the models.

SLIDE 14
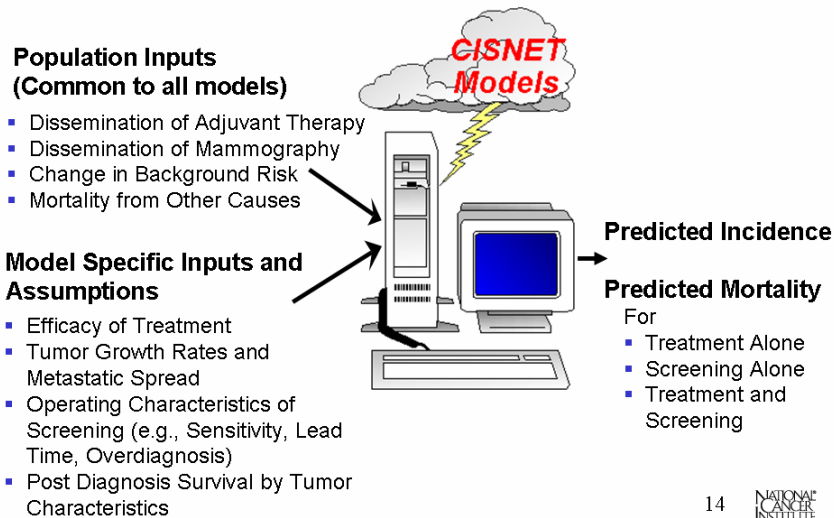
SLIDE 14 NOTES: No notes.

SLIDE 15

Colorectal Cancer Base Case Questions

- Graduated approach

- Base Case I – What is the impact of different natural history models on incidence and mortality
  - Hypothetical situation with conditions in 1978 frozen in time with no interventions

- Base Case II – What is the impact of a single screen (FOBT, colonoscopy, flex sig) at age 65 w/wo follow-up surveillance after detection of an adenoma

15   NATIONAL CANCER INSTITUTE

SLIDE 15 NOTES: No notes.

SLIDE 16

Modeling Hypothetical Cohorts vs. Population Cohorts

- Cost effectiveness models usually conducted on hypothetical cohorts

- In some settings population models can add additional insights to C/E issues especially when:
  - Capacity is an issue
  - Costs depend on volume

16

SLIDE 16 NOTES: No notes.