

Lung Cancer Screening With Helical Computed Tomography in Older Adult Smokers

A Decision and Cost-effectiveness Analysis

Parthiv J. Mahadevia, MD, MPH

Lee A. Fleisher, MD

Kevin D. Frick, PhD

John Eng, MD

Steven N. Goodman, MD, PhD

Neil R. Powe, MD, MPH, MBA

CONCERNED SMOKERS AND THEIR physicians are contemplating screening for lung cancer with helical computed tomography (CT).¹⁻⁴ Having an average 5-year survival of 15%,⁵ lung cancer is often diagnosed at advanced clinical stages when treatment is typically noncurative. Therefore, physicians and their patients have long sought screening methods to detect lung cancer at localized stages, when it can be surgically removed and possibly cured.

Past attempts to screen for lung cancer with intensive screening regimens using chest radiographs and sputum cytology found no reduction in lung cancer-specific mortality,⁶⁻¹⁰ raising skepticism about whether lung cancer is amenable to early detection.¹¹ Uncontrolled studies of helical CT screening report detection of many small-sized lung cancers, when surgical resection and theoretical cure is possible.¹²⁻¹⁹ However, these findings are of unclear clinical significance due to the lack of control groups and of long-term morbidity or mortality data in these trials. Some contend that surgical treatment

See also p 357 and Patient Page.

Context Encouraged by direct-to-consumer marketing, smokers and their physicians are contemplating lung cancer screening with a promising but unproven imaging procedure, helical computed tomography (CT).

Objective To estimate the potential benefits, harms, and cost-effectiveness of lung cancer screening with helical CT in various efficacy scenarios.

Design, Setting, and Population Using a computer-simulated model, we compared annual helical CT screening to no screening for hypothetical cohorts of 100 000 current, quitting, and former heavy smokers, aged 60 years, of whom 55% were men. We simulated efficacy by changing the clinical stage distribution of lung cancers so that the screened group would have fewer advanced-stage cancers and more localized-stage cancers than the nonscreened group (ie, a stage shift). Our model incorporated known biases in screening programs such as lead time, length, and overdiagnosis bias.

Main Outcome Measures We measured the benefits of screening by comparing the absolute and relative difference in lung cancer-specific deaths. We measured harms by the number of false-positive invasive tests or surgeries per 100 000 and incremental cost-effectiveness in US dollars per quality-adjusted life-year (QALY) gained.

Results Over a 20-year period, assuming a 50% stage shift, the current heavy smoker cohort had 553 fewer lung cancer deaths (13% lung cancer-specific mortality reduction) and 1186 false-positive invasive procedures per 100 000 persons. The incremental cost-effectiveness for current smokers was \$116 300 per QALY gained. For quitting and former smokers, the incremental cost-effectiveness was \$558 600 and \$2322 700 per QALY gained, respectively. Other than the degree of stage shift, the most influential parameters were adherence to screening, degree of length or overdiagnosis bias in the first year of screening, quality of life of persons with screen-detected localized lung cancers, cost of helical CT, and anxiety about indeterminate nodule diagnoses. In 1-way sensitivity analyses, none of these parameters was sufficient to make screening highly cost-effective for any of the cohorts. In multiway sensitivity analyses, a program screening current smokers was \$42 500 per QALY gained if extremely favorable estimates were used for all of the influential parameters simultaneously.

Conclusion Even if efficacy is eventually proven, screening must overcome multiple additional barriers to be highly cost-effective. Given the current uncertainty of benefits, the harms from invasive testing, and the high costs associated with screening, direct-to-consumer marketing of helical CT is not advisable.

JAMA. 2003;289:313-322

www.jama.com

Author Affiliations: Robert Wood Johnson Clinical Scholars Program (Drs Mahadevia and Powe); Departments of Medicine (Drs Mahadevia and Powe), Anesthesiology (Dr Fleisher), Radiology (Dr Eng), and Oncology (Dr Goodman), Johns Hopkins School of Medicine, Baltimore, Md; Departments of Health Policy and Management (Drs Fleisher, Frick, Goodman, and Powe) and Epidemiology (Drs Goodman and Powe), Johns Hopkins Bloomberg

School of Public Health, Baltimore; Welch Center for Prevention, Epidemiology, and Clinical Research, Johns Hopkins Medical Institutions, Baltimore (Dr Powe); and MEDTAP International, Bethesda, Md (Dr Mahadevia).

Corresponding Author and Reprints: Parthiv J. Mahadevia, MD, MPH, MEDTAP International, 7101 Wisconsin Ave, Bethesda, MD 20814 (e-mail: Mahadevia@medtap.com).

of these small-sized lung cancers should lower mortality,²⁰ whereas others are doubtful of a mortality benefit due to tendency of some lung cancers to micrometastasize even at apparently localized stages.^{18,21} Although large randomized controlled trials evaluating helical CT as a screening test for lung cancer are beginning, they are years from completion.^{22,23}

Meanwhile, direct-to-consumer marketing⁴ and media coverage of the helical CT trials has encouraged demand for lung cancer screening,^{2,3,24} despite the lack of evidence. Early dissemination of a screening technology raises concerns such as consequences of false-positive and false-negative results,⁴ harms from invasive diagnostic tests,²⁵ surgery,¹ and considerable societal costs.²⁶

We studied the potential benefits, harms, and cost-effectiveness of annual helical CT screening under various scenarios of efficacy. Our research aim was to gain insight into the important factors influencing screening efficacy and potential economic and safety consequences of a widely disseminated lung cancer screening program.

METHODS

Study Design

We performed a decision and cost-effectiveness analysis from the societal perspective in which we simulated the clinical paths and health states that persons would traverse having received or not received annual helical CT screening. We considered parameters required for evaluating a screening program, such as the characteristics of the participants (age, smoking status), characteristics of the screening test (sensitivity, specificity), characteristics of the disease (treatment and prognosis, quality of life), and characteristics of the screening program itself (efficacy, biases inherent in screening). Probabilities for these parameters,^{12-15,17,27-34} as well as quality of life measured in the form of utilities^{35,36} and costs,^{3,37-42} were obtained from the published literature and the Surveillance, Epidemiology, and End Results (SEER) national cancer database²⁷ (TABLE 1).

We performed a base-case analysis, which measures cost-effectiveness using the estimate thought to be the most accurate for each parameter in the model. We performed 1-way sensitivity analyses, which evaluate the impact on cost-effectiveness of using more extreme estimates of each parameter. Through sensitivity analyses, model parameters that are influential in determining the cost-effectiveness of screening can be identified. Finally, we performed multiway sensitivity analyses, which vary estimates across more than 1 parameter. These analyses depict economic consequences across a range of favorable and unfavorable circumstances. We created a Markov model to perform these analyses.⁴³

Study Population

Our study population consisted of 100 000 hypothetical 60-year-old heavy smokers (>20 pack-years) who were eligible for lung resection surgery. Fifty-five percent of the cohort were men. We chose this population since they mirror participants in screening trials and are at higher lung cancer risk than the overall population, making them more likely to benefit from screening.

Because smoking cessation might impact the benefit of screening, we performed separate analyses for 3 different cohorts of heavy smokers: persons who were continuing smoking (current smokers), persons who had permanently quit smoking at time of initial screening (quitting smokers), and persons who had permanently quit 5 years prior to screening and survived to age 60 years (former smokers). We modeled annual screening from age 60 to age 80 years (20-year screening duration) with follow-up until age 100 years (40-year time horizon for occurrence of clinical events). We terminated screening at age 80 years since the incidence of lung cancer decreases afterward.²⁸

The Markov Model and Its Transition Probabilities

Nonscreened Cohort. Smokers in the nonscreened group (FIGURE 1) face

yearly transition probabilities of staying alive without apparent lung cancer, developing and dying from lung cancer, or dying from other causes. We obtained these incidence and mortality probabilities from the SEER DevCan program²⁸ (Table 1). We adjusted these measures for the 3 smoking cohorts using estimates of relative risk corresponding to smoking cessation.^{44,45} All probabilities were stratified by 5-year age groups and sex-adjusted to our reference population.

Persons diagnosed with suspicious lesions or symptoms suggestive of lung cancer undergo invasive testing, with its potential harms. Those diagnosed with lung cancer could have localized-stage non-small-cell lung cancer (NSCLC), advanced-stage NSCLC, or small-cell lung cancer (SCLC). Upon entering these disease states, persons undergo various clinical management strategies (surgery vs radiation therapy vs chemotherapy) and have distinct prognoses. Annual survival probabilities for each histological clinical stage were obtained from the SEER program using lung and bronchus cancers that were invasive, microscopically confirmed, actively followed, and with no other primary cancer. We excluded SEER cases from nursing homes and those identified only by autopsy data. We classified SEER cases into localized or advanced clinical stages using the SEER-modified American Joint Committee Classification. Stage IA and IB cases receiving curative resection (segmentectomy, lobectomy, or pneumonectomy) were classified as the localized-stage NSCLC group and all other stages including unresected stage IA through IB cases formed the advanced-stage NSCLC group. We stratified all probabilities by age, histological-clinical staging, and disease duration. We obtained the proportion of mortality attributable to lung cancer vs other causes from SEER cause-of-death data.

Screened Cohort. Smokers undergoing screening have similar risks of developing lung cancer but face a different course of events. Adherent smokers in the screened group receive annual

Table 1. Annual Probabilities, Costs, and Utilities*

	Base-Case Analysis	Sensitivity Analysis		Reference
		Favors Screening	Against Screening	
Lung cancer probabilities				
Proportion of lung cancers that are NSCLC, %	85	85	83	27
Proportion of non-screen NSCLC that are advanced stage, %	84	85	83	27
Mortality, %				
Localized-stage NSCLC†‡	10	5	10	27
Advanced-stage NSCLC†‡	63	63	63	27
SCLC†‡	61	61	61	27
Persons without lung cancer†§	1.27	1.27	1.27	28
Lung cancer incidence, current smokers, %†§	0.43	0.43	0.43	28
Length bias, % of incidence				
Baseline screen	200	150	350	Estimate
Repeat screen	0	0	10	Estimate
Harms from treatment, %				
Death from invasive testing	0.02	0.01	0.04	29-32
Frequency of complications after invasive testing	14	8	20	29
30-day mortality after surgical resection	4.5	3	6	34
Costs, \$				
Screening				
Helical CT	300	150	450	3
Follow-up diagnostic CT	429	300	500	37
Antibiotic course	158	126	190	38
Opportunity costs or cost for travel time	14	8	20	Estimate
Cancer care				
Localized NSCLC, surviving first year	43 900	35 120	52 680	39
Localized NSCLC, dying in first year	66 500	53 200	79 800	39
Localized NSCLC, ongoing year	4500	3600	5400	39
Localized NSCLC, terminal year	30 400	24 320	36 480	39
Other				
Informal caregiving, localized NSCLC	1800	1440	2160	40
Informal caregiving, advanced NSCLC	6200	7440	4960	40
Diagnostic workup	2300	1840	2760	39
Specialty visit, new	168	84	252	37
Level 4 visit, established	79	60	100	37
Chest radiograph	37	30	44	37
Surgery	42 800	34 240	51 360	41
Diagnostic complications	2700	1350	4050	42
Fine-needle aspiration biopsy	360	300	420	37

Quality-of-Life Weights: Utility or Disutility From Screening, Treatment, and Health States

	Base-Case Analysis	Favors Screening	Against Screening	Time Frame	Source
Utility					
Localized NSCLC nonscreening group	0.73	0.69	0.83	Life	35
Localized NSCLC screening group	0.83	0.88	0.69	Life	35
Advanced NSCLC or SCLC	0.66	0.30	0.76	Life	35
Disutility					
Unclear diagnosis—indeterminate nodule	-0.03	0	-0.06	120 days	36
Invasive diagnostic testing	-0.03	0	-0.05	10 days	36
Surgery/postoperative period	-0.18	-0.05	-0.23	30 days	36
Radiation and/or chemotherapy	-0.14	-0.19	-0.03	90 days	36

Abbreviations: CT, computed tomography; NSCLC, non-small-cell lung cancer; SCLC, small-cell lung cancer.

*The model assumes a 50% stage shift. Estimates for screening program sensitivity, specificity, and nonadherence are shown in Table 2.

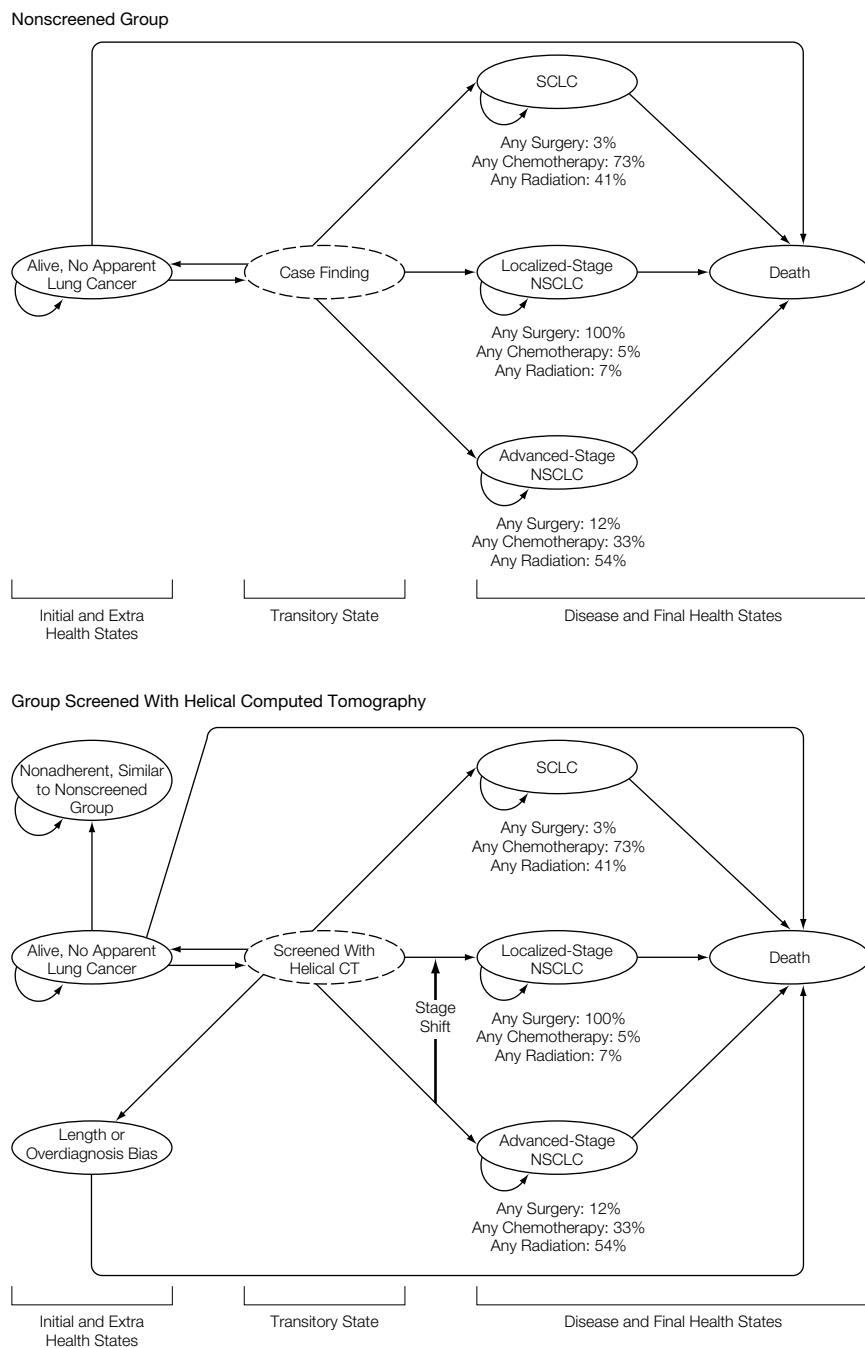
†First-year probability only.

‡Annual mortality probability conditioned on age at the time of diagnosis, stage at diagnosis, and disease duration.

§Annual probability conditioned on age and duration of smoking cessation.

||Costs for advanced-stage disease (not shown) were obtained from the same source as those for localized-stage disease.

Figure 1. Markov Model of Annual Lung Cancer Screening With Helical CT



Both the nonscreened and screened groups start at the “alive, no apparent lung cancer” state. Lung cancer cases are identified by symptomatic or incidental case finding (nonscreened group) or screening with helical computed tomography (CT) (screened group). Cases transition to disease states after undergoing a workup state (broken-lined circles). A certain proportion will be diagnosed as having non-lung cancerous nodules and return to the “alive, no apparent lung cancer” state. Those correctly identified with lung cancer progress to disease states and receive treatment specific to their clinical situation. Treatments in the first year after diagnosis are shown below each disease state (from the Surveillance, Epidemiology, and End Results database²⁷ and the National Cancer Database Report.⁵² 1995). Based on age, histology, clinical stage, and duration of diagnosis, a certain proportion die annually. Screening efficacy is derived from a stage shift (percent decrease in advanced-stage non-small-cell lung cancer [NSCLC], with a corresponding percent increase in localized-stage NSCLC). SCLC indicates small-cell lung cancer.

helical CT testing and based on its accuracy may be diagnosed with indeterminate nodules that require additional monitoring and possible invasive testing. We obtained probabilities for the frequency of indeterminate nodules, screening sensitivity and specificity, and annual adherence from helical CT trials that have reported baseline and annual repeat screening data (TABLE 2).^{12-15,17}

The timing of cancer diagnoses, the numbers of diagnoses, and the types of cancers among screened individuals can be different than those who are not screened. Screened persons could be diagnosed with lung cancers that do not cause clinical disease (overdiagnosis bias) or are extremely slow-growing and have long latency periods (length bias).⁴⁶ Many of these individuals might remain undiagnosed under no-screening conditions. Evidence for these biases has been described in past screening trials.^{46,47} We estimated the degree of length bias and overdiagnosis bias by comparing trial-specific observed diagnosis rates with expected diagnosis rates (Table 1). Expected diagnosis rates were calculated using incidence probabilities weighted for average trial-specific age, sex, and smoking composition.

Screening also creates a lead time (advancement in the date of diagnosis), in which the timing of diagnostic testing, improvements in health, and costs occur earlier than under no-screening conditions. We incorporated an average 1-year lead time for screening-detected lung cancers, extrapolated from published lead times for lung cancer screening using chest radiographs⁴⁸ and tumor doubling times.¹³

Screening-detected cancers also have a different histological distribution. Cancers detected by helical CT (ie, true-positive results) are usually peripherally based cancers, while those missed by helical CT (ie, false-negative results) are often hidden in endobronchial locations.¹³ Endobronchial tumors are more frequently due to squamous cell carcinoma, while peripheral tumors tend to represent adenocarcinoma. In a more complex model, we

Table 2. Helical CT Screening Studies: Participants, Accuracy, and Adherence Outcomes

	Trial				Weighted Average or Totals
	Henschke et al, ¹² 1999 Henschke et al, ¹³ 2001	Swensen et al, ¹⁵ 2002	Sone et al, ¹⁴ 2001	Sobue et al, ¹⁷ 2002	
Location	United States	United States	Japan	Japan	
Design	Prospective uncontrolled cohort study	Prospective uncontrolled cohort study	Prospective uncontrolled cohort study	Prospective uncontrolled cohort study	
Years	1993-1998	1999-2002	1996-1998	1993-1998	
Recommended screening interval	1 Year	1 Year	1 Year	6 Months	
Participant demographics					
Men, %	54	52	54	88	
Age, mean or median, y	67	60	≥65	60	
Smoking status	44 pack-years	CS (61%), FS (39%)	ES (46%), NS (54%)	CS (62%), FS (25%), NS (14%)	
Helical CT scan parameters (all protocols used 120-140 kVp and between 25-50 mA)	Single-detector 10-mm slice thickness	Multidetector 5-mm slice thickness	Single-detector 10-mm slice thickness	Single-detector 10-mm slice thickness	
Baseline Screening Year					
Participants screened, No.	1000	1520	5483	1611	9614
Screening tests performed, No.	1000	1520	5483	1611	9614
Outcome measure					
New indeterminate lung nodules detected	233	782	676	192	1883
Rate of new indeterminate nodules per 1000 screenings	233	514	123	119	196
Lung cancers detected (true-positive)	31	21	22	13	87
Lung cancers missed (false-negative)	2	2	18	1	23
Lung nodules, noncancerous (false-positive)	200	759	636	178	1773
Negative helical CT scans (true-negative)*	767	738	4807	1419	7731
Benign biopsies or surgeries (harm)	3	4	7	8	22
Screening program performance					
Sensitivity, %†	94	91	55	93	93
Specificity, %	79	49	88	89	81
Rate of lung cancers detected per 1000 screenings	31.0	13.8	4.0	8.1	9.0
Rate of benign biopsies/surgeries per 1000 screenings	3.0	2.6	1.3	5.0	2.3
Annual Repeat Screening Years					
Participants screened, No.	841	1464	4425	1180	11 788
Screening tests performed, No.	1184	1464	8303	7891	18 842
Outcome measure					
New indeterminate lung nodules detected	30	191	518	770	1509
Rate of new indeterminate nodules per 1000 screenings	25	130	62	98	80
Lung cancers detected (true-positive)	7	2	25	19	53
Lung cancers missed (false-negative)	ID	ID	5	3	8
Lung nodules, noncancerous (false-positive)	ID	ID	488	748	1236
Negative helical CT scans (true-negative results)*	ID	ID	7785	7121	14 906
Benign biopsies or surgeries (harm)	2	3	9	27	41
Annual Repeat Screening Years					
Screening program performance					
Sensitivity, %	ID	ID	83	86	85
Specificity, %	ID	ID	94	90	92
Rate of lung cancers detected per 1000 screenings	5.9	1.4	3.0	2.4	2.8
Rate of benign biopsies/surgeries per 1000 screenings	1.7	2.0	1.1	3.4	2.1
Annual nonadherence, %‡	12	3	18	25	6.5

Abbreviations: CS, current smoker; CT, computed tomography; ES, ever smoker; FS, former smoker; ID, insufficient data (due to lack of long-term follow-up results); kVp, kilovolt peak; NS, never smoker.

*False-negatives were lung cancers identified by sputum cytology only, missed in prior screening or interval cases. All false-negatives have been reassigned to prior year's screening.

†Sensitivity calculation omits Sone et al¹⁴ due to insufficient triage protocol in baseline screening year.

‡We used an average nonadherence rate of 6.5% from US studies only (range, 3%-12%). Japanese studies used less-organized population-based screening, resulting in poor adherence.

adjusted lung cancer survival estimates to account for this histological predilection. However, because this adjustment did not significantly impact our results, the results shown are from a simpler model that did not incorporate this phenomenon.

Finally, screening should alter the clinical-stage distribution of cancers such that advanced cancers are found earlier in localized stages. This is frequently referred to as a downward stage shift. In our base-case scenario, we assumed a 50% stage shift. For example, if the stage distribution for the non-screened group is 80% advanced and 20% localized, then a 50% stage shift would result in a 40% advanced- and 60% localized-stage distribution for the screened group.

In sensitivity analyses we varied the degree of stage shift and also examined the impact of a pseudo-stage shift. For lung cancer, a pseudo-stage shift can occur if the advanced-stage cancers that have been “shifted” into localized stages are biologically aggressive and carry a prognosis closer to naturally occurring advanced-stage disease than to localized-stage disease.

Model Assumptions. We made several major assumptions for this study: (1) Only NSCLC underwent a stage shift; SCLC did not experience a stage shift given its early metastatic nature and low probability of cure. (2) Within histological and clinical stage categories, the clinical management, time costs of treatments, treatment risks, and subsequent prognoses were on average similar between screened and non-screened lung cancer cases. This assumption was removed for the pseudo-stage-shift scenario. (3) The incidence of lung cancers in the screened group did not increase due to the radiation exposure from CT scans. (4) Persons who became nonadherent to screening remained nonadherent. Therefore, partial adherence or intermittent screening was not modeled. (5) Lung cancer cases due to length bias and overdiagnosis bias received the same costs, workup, and therapy as all other lung cancer cases but had mortality prob-

abilities similar to smokers without lung cancer. (6) The quality of life for persons with screening-detected localized NSCLC was higher than non-screening-detected localized NSCLC since screening-detected cases typically do not have symptoms at the time of diagnosis.

Costs

The costs of screening include the facility and professional fees associated with helical CT screening, the monitoring of indeterminate nodules, and the opportunity costs to get screening (Table 1). Cost of cancer care stratified by clinical stage and duration of illness was obtained from a study using a SEER registry linked to a health maintenance organization database with chart reviews.³⁹ We obtained costs for specific diagnostic tests and physician visits from the American Medical Association’s 2001 National Physician Fee Schedules Relative Value Scale.³⁷ We obtained the costs for informal caregiving,⁴⁰ treatment, and its complications from the literature.^{29,33,41,42} All costs were calibrated to 2001 US dollars using the medical component of the consumer price index. Both costs and life-years were discounted 3% annually.

Quality of Life

We obtained quality-of-life measurements for our health states from a study of patients with NSCLC³⁵ using results from the EuroQoL multiattribute utility scale⁴⁹ (Table 1). We used the median utility scores for patients without and with metastasis as proxies for the localized and advanced NSCLC/SCLC groups, respectively. Since duration of disease did not statistically alter utilities, we used these scores as the average utility for the remainder of case lifetime.

From a systematic review of cost-utility assessments in oncology, we obtained the decrements in quality of life or disutilities for anxiety and discomfort from having an unclear noncalcified nodule, invasive testing, surgery, and radiation/chemotherapy.³⁶

Outcome Measures

We measured the benefits of screening by comparing the absolute and relative difference in lung cancer-specific deaths. We estimated the number of screening tests performed and harms from invasive tests or surgery and deaths from lung cancer for every 100 000 persons screened. Costs were measured in US dollars and incremental effectiveness in quality-adjusted life-years (QALYs) gained.

Sensitivity Analyses

To explore the benefits, harms, and cost-effectiveness of screening under different efficacy assumptions we varied the degree of stage shift and examined the impact of a pseudo-stage shift. To examine the ideal age range to start screening, we varied age at first screening from ages 45 to 80 years. To examine the time to reach cost-effectiveness, we varied the length of follow-up from 1 to 40 years.

After performing 1-way sensitivity analyses to isolate influential parameters, we examined the impact of simultaneously using highly favorable and unfavorable estimates for the most influential model parameters other than stage shift.

RESULTS

Base-Case Analysis

For current smokers in the base-case scenario, there were 462 352 screening examinations over a 20-year period. Using a 50% stage shift assumption, there were 4168 lung cancer deaths per 100 000 nonscreened persons compared with 3615 lung cancer deaths per 100 000 screened persons, resulting in an absolute mortality reduction of 553 deaths or a 13% relative mortality reduction (TABLE 3). However, there were also 1186 invasive tests or surgeries for benign lesions in the screened group. The incremental cost-effectiveness of screening was \$116 300 per QALY gained.

The incremental costs consumed by screening (TABLE 4) were between \$4300 and \$4600 for all smoking cohorts. However, the incremental effec-

tiveness of screening decreased with lower-risk cohorts; there was a 0.039-QALY gain among current smokers, a 0.008-QALY gain among quitting smokers, and a 0.002-QALY gain among former smokers. The incremental cost-effectiveness for screening quitting smokers was \$558 600 per QALY gained and for former smokers was

\$2322 700 per QALY. Thus, annual screening with helical CT became progressively less cost-effective among smokers with longer duration of smoking cessation.

Sensitivity Analyses

Degree of Stage Shift. For current smokers, we varied the degree of stage

shift from 0% to 100%. The stage shift required for screening to cost less than \$50 000 per QALY was 91% for current smokers. Screening was dominated by no screening if less than a 23% stage shift occurred. For quitting and former smokers, a 100% stage shift was not sufficient by itself to reach \$50 000 per QALY gained.

Table 3. Multiway Sensitivity Analysis of Lung Cancer Screening With Helical CT Among Current Smokers Only

Analysis	Outcomes Measured	Change From No Screening
Base-Case Scenario		
Assumes a 50% stage shift*	Change in lung cancer deaths, per 100 000	-553
Annual nonadherence: 6.5%	Change in lung cancer mortality, %	-13
Cost of helical CT: \$300	Number harmed, per 100 000	1186
Utility of screening-detected NSCLC: 0.83	Incremental cost-effectiveness ratio, \$	116 300
Degree of length bias and overdiagnosis bias: 200% of nonscreened cancers in prevalence-year screening		
Favorable Estimate Scenario		
Assumes 50% stage shift and makes following changes to base-case scenario:	Change in lung cancer deaths, per 100 000	-900
Annual nonadherence decreased to 3%	Change in lung cancer mortality, %	-16
Cost of helical CT decreased to \$150	Number harmed, per 100 000	1520
Utility of screening-detected localized NSCLC increased to 0.88	Incremental cost-effectiveness ratio, \$	42 500
Degree of length bias and overdiagnosis bias decreased to 150% of nonscreened cancers in prevalence-year screening		
Anxiety from unclear diagnosis decreased to 0		
Unfavorable Estimate Scenario		
Assumes 50% stage shift and makes following changes to base-case scenario:	Change in lung cancer deaths, per 100 000	-119
Annual nonadherence increased to 12%	Change in lung cancer mortality, %	-4
Cost of helical CT increased to \$450	Number harmed, per 100 000	993
Utility of screening-detected localized NSCLC decreased to 0.73	Incremental cost-effectiveness ratio	Dominated†
Degree of length bias and overdiagnosis bias increased to 250% of nonscreened cancers in prevalence-year screening		
Anxiety from unclear diagnosis increased to -0.06		
Pseudo-Stage-Shift Scenario		
Assumes that 50% of advanced-stage cancers are labeled as localized and undergo curative surgery; however, these cancers behave aggressively, micrometastasize, and carry advanced-stage prognosis	Change in lung cancer deaths, per 100 000	6
	Change in lung cancer mortality, %	-0.1
	Number harmed, per 100 000	1186
	Incremental cost-effectiveness ratio	Dominated†

Abbreviations: CT, computed tomography; NSCLC, non-small-cell lung cancer.

*Reduction in advanced-stage cancers (with proportionate increase in localized stage cancers). The shifted advanced cancers are surgically resected and have a prognosis similar to localized lung cancers.

†No screening is more preferable than screening (screening has higher costs and lower effectiveness).

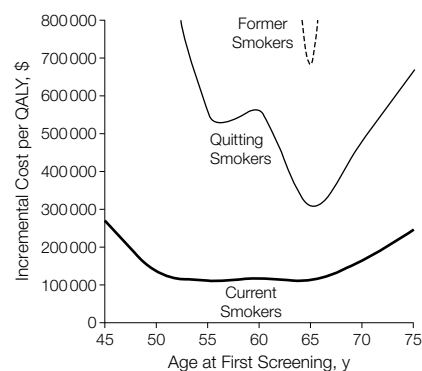
Table 4. Base-Case Results Comparing Helical CT Screening With No Screening for Current, Quitting, and Former Smokers*

	Current Smokers		Quitting Smokers		Former Smokers	
	Nonscreened	Screened	Nonscreened	Screened	Nonscreened	Screened
Cost, \$						
Average	5400	9900	3600	8100	2800	7100
Incremental		4600		4500		4300
QALYs						
Average	12.96	13.00	14.38	14.39	14.77	14.77
Incremental		0.039		0.008		0.002
Incremental cost per QALY, \$		116 300		558 600		2 322 700

Abbreviations: CT, computed tomography; QALY, quality-adjusted life-year.

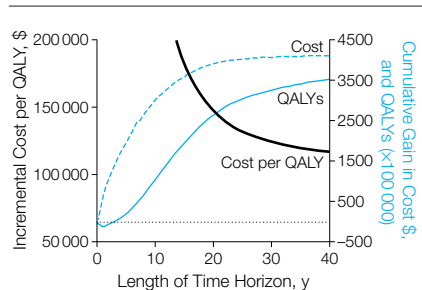
*The incremental cost and incremental QALYs of screening are the difference in the average values of the nonscreened group from the screened group. The incremental cost-effectiveness ratio is the incremental costs over incremental QALYs. Numbers shown have been rounded and therefore incremental cost-effectiveness ratios shown are more exact than manual calculations.

Figure 2. Effect of Varying Age at First Screening vs Incremental Cost-effectiveness of Helical Computed Tomography Screening for Current, Quitting, and Former Smokers



For all cohorts, screening is most cost-effective prior to the peak incidence of lung cancer (ages 67-72 years). Dominated indicates that no screening is preferable to screening, since screening has higher costs and lower effectiveness. QALY indicates quality-adjusted life-year.

Figure 3. Effect of Varying the Duration of Follow-up on Cumulative Change in Cost and QALYs and Incremental Cost-effectiveness of Helical CT Screening Among Current Smokers



There is a net loss in quality-adjusted life-years (QALYs) during the first 2 years of screening from the increased workups, quality-of-life impairments, and excess cases diagnosed by computed tomography (CT) screening. After 19 years the incremental cost-effectiveness ratio drops below \$150 000 per QALY.

Age at First Screening. Cost-effectiveness varied with age at first screening in a U-shaped fashion (FIGURE 2). For current smokers, screening was most cost-effective when started between ages 55 and 65 years. For each

smoking cohort, screening was most cost-effective when started before the ages of peak lung cancer incidence (ages 67-72 years). Screening was dominated by no screening for persons older than age 77 years among current smokers, older than age 75 among quitting smokers, and older than age 74 for former smokers. Screening was also dominated for persons younger than age 48 for quitting smokers and persons younger than age 50 for former smokers.

Length of Follow-up. To assess how quickly the costs, QALYs, and cost-effectiveness of screening would accrue, we varied the time horizon or duration of follow-up. We evaluated the cumulative change in costs, QALYs, and incremental cost-effectiveness for current smokers (FIGURE 3). For the first 2 years, there was a net loss in QALYs due to the harms and disutilities from testing and treatment. Gains in QALYs from screening and treatment did not manifest until the third year of follow-up. There were significant up-front costs for screening, with the largest increase occurring in the first year. The incremental cost-effectiveness ratio approached \$120 000 per QALY gained after 19 years of follow-up.

Influential Model Parameters and Multiway Sensitivity Analysis. We performed sensitivity analysis on the remaining model parameters. Parameters that changed incremental cost-effectiveness by more than 50% of base case were classified as highly influential. In addition to stage shift, nonadherence with screening, length bias, or overdiagnosis bias in the prevalence year of screening, quality of life for screening-detected localized NSCLC, costs for helical CT, and anxiety from expectant management of indeterminate nodules were highly influential.

In a favorable estimate scenario (Table 3) using lower probabilities for annual nonadherence (3% instead of 6.5%), lower estimates for length bias and overdiagnosis bias (150% of incidence instead of 200%), better quality of life for localized-stage lung cancers (0.88 instead of 0.83), lower costs per case for helical CT screening (\$150 in-

stead of \$300), and no anxiety from indeterminate nodules, screening led to 900 fewer lung cancer deaths (16% mortality reduction), 1520 false-positive invasive procedures, and an incremental cost-effectiveness of \$42 500 per QALY gained for current smokers. Using a similar scenario for the other risk cohorts resulted in an incremental cost-effectiveness ratio of \$75 300 per QALY gained for quitting smokers and \$94 400 per QALY gained for former smokers.

Under an unfavorable estimate scenario for current smokers (Table 3 shows point estimate changes), there were 119 fewer lung cancer deaths in the screened group (4% relative mortality reduction) and 993 individuals harmed. The incremental cost-effectiveness of screening was dominated by no screening for all smoking cohorts. Increasing nonadherence or the degree of length bias and overdiagnosis bias alone caused the cost-effectiveness of screening to be dominated by no screening.

If screening resulted in a pseudo-stage shift whereby shifted cancers had high potential for micrometastasis and did not carry the typical favorable prognosis of localized NSCLC, screening resulted in 6 fewer lung cancer deaths (0.1% relative mortality reduction) and 1186 false-positive invasive procedures, and its cost-effectiveness was dominated by no screening.

COMMENT

Substantial debate about the use of imaging modalities such as helical CT for lung cancer screening arises from the lack of definitive evidence about its efficacy, high societal costs, and the potential for harm. In this study, assuming a sizable stage shift and incorporating relevant considerations for evaluating a screening program, we found that lung cancer screening with helical CT is unlikely to be highly cost-effective for most heavy smokers.

While a scenario can be envisioned under which screening is cost-effective, the likelihood of such a scenario manifesting itself is question-

able. Multiple barriers associated with screening must be overcome. These barriers include nonadherence to screening, high screening costs, poor quality of life after diagnosis, excess testing and procedures, excess diagnoses from length bias and overdiagnosis bias, and having to wait many years to achieve the economic and health benefit from screening.

The total societal cost for an annual helical CT screening program of at-risk ever-smokers is very high. An estimated 50 million men and women in the United States are ever-smokers between the ages of 45 and 75 years.²⁶ If 50% of this group received periodic annual screening, the program costs are approximately \$115 billion (discounted) based on our study estimates.

Our findings raise concern regarding direct-to-consumer marketing. Frequently, screening advertisements found on the Internet and in newspapers tout the benefits of early detection, do not clarify who is at high risk for lung cancer, and do not mention the risks of finding indeterminate nodules. If the aforementioned 25 million individuals underwent 1-time helical CT screening only, we estimate the number of individuals having indeterminate nodules diagnosed to be approximately 5 million. Even among high-risk groups, most nodules are non-cancerous. If screening were offered to groups at lower risk, the frequency of unnecessary invasive testing or surgeries could increase because pulmonary nodules are more likely to be benign in groups at lower risk.

For our simulated cohort, screening resulted in more than 1000 false-positive cases undergoing invasive testing and/or surgery for benign lesions. Our estimate is from the helical CT clinical trials, which used detailed radiographic monitoring protocols to prevent unnecessary invasive procedures. Widespread dissemination of screening could result in greater harm if quality-of-care standards are lower in the general population. These findings warrant caution in direct-to-consumer marketing. For those persons enquiring

about helical CT screening, appropriate informed consent should discuss the risk of harm and weigh the likelihood of benefiting from screening.

Our finding that the risk profile of the screened population markedly affects cost-effectiveness has important implications. Groups defined by simple demographic factors, such as age or smoking status, may not be at high enough risk to make routine, population-based screening cost-effective. Identification of groups at even higher risk for screening is likely to be the most cost-effective strategy, although this will limit screening to a smaller population. Clinical factors, such as a history of asbestos exposure among smokers or the presence of obstructive airway disease, may identify groups at higher risk. However, many individuals with these characteristics may not qualify for surgery due to the severity of comorbid conditions, making screening less beneficial. Identification of high-risk groups through biomarkers may be promising. Recently, Palmisano and colleagues⁵⁰ found hypermethylation of cancer-fighting genes in the sputum cells of smokers. Examination of sputum specimens several years before cancer developed revealed that hypermethylation antedated lung cancer in each of 21 persons who eventually developed lung cancer. This specific biomarker or others developed in the future could identify higher-risk groups for whom subsequent screening with an imaging modality might be useful.

In this study, we make several assumptions that bias our model in favor of screening. For example, we assumed that the radiation from the scans would not increase risk of lung cancer. We assigned higher quality-of-life scores for screening-detected localized-stage NSCLC compared with non-screened cases. By favoring screening efficacy, these assumptions make our conclusion about the cost-ineffectiveness of screening more conservative.

There are several limitations to this study. First, our model did not include benefits, harms, and costs for incidental diagnoses from screening. For ex-

ample, researchers at the Mayo Clinic CT screening program reported incidental findings among 210 participants, including 51 abdominal aortic aneurysms, 33 indeterminate renal masses, 35 adrenal masses, and 24 renal calculi.¹⁵ The benefits, harms, and costs for diagnosis and management of these conditions, although not firmly established, could be substantial. However, we doubt that early detection of these incidental findings would improve life expectancy enough to justify lung cancer screening if screening for lung cancer is not by itself cost-effective. Second, this study is a model of clinical practice rather than an examination of actual clinical practice. However, our model incorporates the wealth of epidemiologic data on lung cancer and does take into account the biases inherent in screening programs. We also performed extensive sensitivity analyses that examined multiple efficacy scenarios. Weighing the risks and benefits of these scenarios can assist policy makers in making screening recommendations and setting guidelines.⁵¹ Such analyses may be particularly helpful when definitive clinical trial data will not be available for years.²³ Third, our study did not incorporate costs associated with disability or lost productivity; however, productivity costs may be a less-prominent parameter in our study given the older age and the heavy comorbidities of the target population. Finally, future advancements in lung cancer diagnosis and treatment could make our results out of date.

In summary, we conclude that lung cancer screening with helical CT is unlikely to be highly cost-effective without substantial reductions in mortality, high rates of adherence, lower rates of overdiagnosis, and lower costs per screening test. Given the uncertainty of efficacy, the possibilities of harm, and the high costs associated with screening, direct-to-consumer marketing of helical CT screening is not advisable.

Author Contributions: *Study concept and design:* Mahadevia, Fleisher, Eng, Powe.
Acquisition of data: Mahadevia.
Analysis and interpretation of data: Mahadevia, Fleisher, Frick, Eng, Goodman, Powe.
Drafting of the manuscript: Mahadevia, Powe.

Critical revision of the manuscript for important intellectual content: Mahadevia, Fleisher, Frick, Eng, Goodman, Powe.

Statistical expertise: Mahadevia, Goodman.

Obtained funding: Powe.

Administrative, technical, or material support: Mahadevia, Frick, Eng, Powe.

Study supervision: Fleisher, Goodman, Powe.

Funding/Support: Dr Mahadevia did this work as a Robert Wood Johnson Clinical Scholar funded by the Robert Wood Johnson Foundation, Princeton, NJ.

REFERENCES

- Lerner B. Screening for cancer: a downside. *New York Times*. March 5, 2002:D6.
- Kolata G. Lung cancer test is much in demand, but benefit is murky. *New York Times*. June 21, 2001:A1-A14.
- Brice J. Low-dose CT: the promise and the paradox of lung cancer screening. *Diagn Imaging (San Franc)*. 2000;22(5):44-53.
- Lee TH, Brennan TA. Direct-to-consumer marketing of high-technology screening tests. *N Engl J Med*. 2002;346:529-531.
- Travis WD, Travis LB, Devesa SS. Lung cancer. *Cancer*. 1995;75(suppl 1):191-202.
- Fontana RS, Sanderson DR, Woolner LB, Taylor WF, Miller WE, Muhm JR. Lung cancer screening: the Mayo program. *J Occup Med*. 1986;28:746-750.
- Tockman MS. Survival and mortality from lung cancer in a screened population: the Johns Hopkins study. *Chest*. 1986;89(suppl 4):324S-326S.
- Melamed MR, Flehinger BJ, Zaman MB. Impact of early detection on the clinical course of lung cancer. *Surg Clin North Am*. 1987;67:909-924.
- Kubik A, Parkin DM, Khlal M, Erban J, Polak J, Adamec M. Lack of benefit from semi-annual screening for cancer of the lung: follow-up report of a randomized controlled trial on a population of high-risk males in Czechoslovakia. *Int J Cancer*. 1990;45:26-33.
- Fontana RS, Sanderson DR, Woolner LB, et al. Screening for lung cancer: a critique of the Mayo Lung Project. *Cancer*. 1991;67(suppl 4):1155-1164.
- Hulka BS. Screening for cancer: lessons learned. *J Occup Med*. 1986;28:687-691.
- Henschke CI, McCauley DI, Yankelevitz DF, et al. Early Lung Cancer Action Project: overall design and findings from baseline screening. *Lancet*. 1999;354:99-105.
- Henschke CI, Naidich DP, Yankelevitz DF, et al. Early lung cancer action project: initial findings on repeat screenings. *Cancer*. 2001;92:153-159.
- Sone S, Li F, Yang ZG, et al. Results of three-year mass screening programme for lung cancer using mobile low-dose spiral computed tomography scanner. *Br J Cancer*. 2001;84:25-32.
- Swensen SJ, Jett JR, Sloan JA, et al. Screening for lung cancer with low-dose spiral computed tomography. *Am J Respir Crit Care Med*. 2002;165:508-513.
- Sone S, Takahima S, Li F, et al. Mass screening for lung cancer with mobile spiral computed tomography scanner. *Lancet*. 1998;351:1242-1245.
- Sobue T, Moriyama N, Kaneko M, et al. Screening for lung cancer with low-dose helical computed tomography: anti-lung cancer association project. *J Clin Oncol*. 2002;20:911-920.
- Patz EF Jr, Goodman PC, Bepler G. Screening for lung cancer. *N Engl J Med*. 2000;343:1627-1633.
- Kaneko M, Eguchi K, Ohmatsu H, et al. Peripheral lung cancer: screening and detection with low-dose spiral CT versus radiography. *Radiology*. 1996;201:798-802.
- Henschke CI, Yankelevitz DF. CT screening for lung cancer. *Radiol Clin North Am*. 2000;38:487-495.
- Patz EF Jr, Rossi S, Harpole DH Jr, Herndon JE, Goodman PC. Correlation of tumor size and survival in patients with stage IA non-small cell lung cancer. *Chest*. 2000;117:1568-1571.
- National Cancer Institute and the American Cancer Society. Early Lung Cancer Screening Workshop. March 7-8, 2001; Rockville, Md.
- National Institutes of Health, National Cancer Institute. Fifty thousand current and former smokers needed for National Lung Screening Trial [press release]. September 18, 2002. Available at: <http://www.nih.gov/news/pr/sep2002/nci-19a.htm>. Accessibility verified December 18, 2002.
- Spiral CT screens raise scientific, economic issues. *Oncol News Int*. 2000;9:1-30.
- Kolata G. Questions grow over usefulness of some routine cancer tests. *New York Times*. December 30, 2001:A1-A16.
- Grannis F Jr. Lung cancer screening: who will pick up the tab? *Chest*. 2002;121:1388-1390.
- National Cancer Institute. *Surveillance, Epidemiology, and End Results (SEER) Program Public-Use Data (1973-1998)* [database on CD-ROM]. Bethesda, Md: National Cancer Institute; 2001.
- DevCan: Probability of Developing or Dying of Cancer* [computer program]. Bethesda, Md: National Cancer Institute; 2001.
- DeCamp MM Jr, Jaklitsch MT, Mentzer SJ, Harpole DH Jr, Sugarbaker DJ. The safety and versatility of video-thoracoscopy: a prospective analysis of 895 consecutive cases. *J Am Coll Surg*. 1995;181:113-120.
- Jaklitsch MT, DeCamp MM Jr, Liptay MJ, et al. Video-assisted thoracic surgery in the elderly: a review of 307 cases. *Chest*. 1996;110:751-758.
- Celik M, Halezeroglu S, Senol C, et al. Video-assisted thoracoscopic surgery: experience with 341 cases. *Eur J Cardiothorac Surg*. 1998;14:113-116.
- Shah SS, Tsang V, Goldstraw P. Open lung biopsy: a safe, reliable, and accurate method for diagnosis in diffuse lung disease. *Respiration*. 1992;59:243-246.
- Sagel SS, Ferguson TB, Forrest JV, Roper CL, Welton CS, Clark RE. Percutaneous transthoracic aspiration needle biopsy. *Ann Thorac Surg*. 1978;26:399-405.
- Bach PB, Cramer LD, Schrag D, Downey RJ, Gelband SE, Begg CB. The influence of hospital volume on survival after resection for lung cancer. *N Engl J Med*. 2001;345:181-188.
- Trippoli S, Vaiani M, Lucioni C, Messori A. Quality of life and utility in patients with non-small cell lung cancer. *Pharmacoeconomics*. 2001;19:855-863.
- Earle CC, Chapman RH, Baker CS, et al. Systematic overview of cost-utility assessments in oncology. *J Clin Oncol*. 2000;18:3302-3317.
- American Medical Association. *1998 National Physician Fee Schedule Relative Value File*. Chicago, Ill: American Medical Association; 1998.
- Thomson Healthcare. *Red Book: Drug Topics 2001*. Montvale, NJ: Medical Economics Co Inc; 2001.
- Fireman BH, Quesenberry CP, Somkin CP, et al. Cost of care for cancer in a health maintenance organization. *Health Care Financ Rev*. 1997;18:51-76.
- National Alliance for Caregiving (NAC), American Association of Retired Persons (AARP). *Caregiving in the United States: Findings From a National Survey*. Bethesda, Md and Washington, DC: NAC/AARP; 2001.
- Hillner BE, Smith TJ. Cost-effectiveness analysis of three regimens using vinorelbine (Navelbine) for non-small cell lung cancer. *Semin Oncol*. 1996;23(suppl 2 pt 5):25-30.
- Gurley MB, Richli WR, Waugh KA. Outpatient management of pneumothorax after fine-needle aspiration: economic advantages for the hospital and patient. *Radiology*. 1998;209:717-722.
- Data Pro [computer program]. Version 6.0. Boston, Mass: TreeAge Inc; 2002.
- Burns DM. Primary prevention, smoking, and smoking cessation: implications for future trends in lung cancer prevention. *Cancer*. 2000;89(suppl 11):2506-2509.
- Peto R, Darby S, Deo H, Silcocks P, Whitley E, Doll R. Smoking, smoking cessation, and lung cancer in the UK since 1950: combination of national statistics with two case-control studies. *BMJ*. 2000;321:323-329.
- Eddy DM. Screening for lung cancer. *Ann Intern Med*. 1990;112:73-74.
- Melamed MR, Flehinger BJ, Zaman MB, Heelan RT, Perchick WA, Martini N. Screening for early lung cancer: results of the Memorial Sloan-Kettering study in New York. *Chest*. 1984;86:44-53.
- Walter SD, Kubik A, Parkin DM, Reissigova J, Adamec M, Khlal M. The natural history of lung cancer estimated from the results of a randomized trial of screening. *Cancer Causes Control*. 1992;3:115-123.
- Dolan P. Modeling valuations for EuroQoL health states. *Med Care*. 1997;35:1095-1108.
- Palmisano WA, Divine KK, Saccomanno G, et al. Predicting lung cancer by detecting aberrant promoter methylation in sputum. *Cancer Res*. 2000;60:5954-5958.
- Harris RP, Helfand M, Woolf SH, et al. Current methods of the US Preventive Services Task Force: a review of the process. *Am J Prev Med*. 2001;20(suppl 3):21-35.
- Fry WA, Phillips JL, Menck HR. Ten-year survey of lung cancer treatment and survival in hospitals in the United States: a national cancer data base report. *Cancer*. 1999;86:1867-1876.