Cost-effectiveness of Whole-Body CT Screening¹

PURPOSE: To make preliminary estimates of the effectiveness (in life-years) and cost-effectiveness (in costs per life-year) of whole-body computed tomographic (CT) screening.

MATERIALS AND METHODS: Costs and effectiveness (in life-years) of onetime whole-body CT screening relative to those of no screening were calculated by using a decision-analytic model. It was assumed that any benefits from screening were due to earlier detection of disease and improvement in survival relative to survival with routine care. Eight conditions were included in the model: ovarian, pancreatic, lung, liver, kidney, and colon cancer; abdominal aortic aneurysm; and coronary artery disease. Costs of the screening examination, follow-up tests, and patient care were estimated. The base-case analysis was performed for a hypothetical cohort of 500 000 self-referred asymptomatic 50-year-old men. For sensitivity analyses, the age and sex of the cohort were varied. Results were expressed in 2001 U.S. dollars per life-year gained.

RESULTS: Compared with routine care, whole-body CT screening provided minimal gains in life expectancy (0.016 6 years or 6 days) at an average additional cost of $2513 per patient, or an incremental cost-effectiveness ratio of $151 000 per life-year gained. Most patients (90.8%) had at least one positive finding, but only 2.0% had disease; work-up in patients with a false-positive result of screening accounted for 32.3% of total costs ($1720 of $5332). Results were sensitive to the prevalence of disease, the effect of screening on stage of disease at diagnosis, the specificity of screening, and the costs of follow-up for false-positive findings.

CONCLUSION: Even with assumptions favorable to whole-body CT, implementation of onetime screening would not be cost-effective compared with currently funded medical interventions; follow-up for false-positive findings would add a substantial financial burden to the health care system.

The appropriateness of whole-body computed tomographic (CT) screening has become the focus of a growing debate. Unfortunately, there are few hard data on which to base an evaluation of the merits of whole-body CT. Proponents to date have relied on anecdotal evidence to support their claims, while skeptics have had few hard facts on which to base their opposition (1,2).

Evaluating screening is fundamentally different from evaluating diagnostic or therapeutic interventions. Because the majority of individuals who undergo screening tests do not have the disease in question, the costs and potential risks are spread across a wider group of individuals. In addition, the costs and potential harms of screening generally occur in the short term, while any benefits are typically realized in the long term.

The most appropriate endpoint for evaluating screening effectiveness is population disease-specific mortality (3). Disease-specific mortality provides an overall estimate of screening effectiveness by giving a combined assessment of the effects of early diagnosis and therapy. Investigators in large randomized trials with long-term follow-up are generally required to estimate the mortality reduction associated with screening. Because of the financial and logistic challenges associated with randomized screening trials, researchers have sought to evaluate screening on the basis of its effect on surrogate outcomes such as stage at diagnosis or patient survival. Unfortunately, endpoints other than disease-specific
mortality have limitations related to lead-time, length, and overdiagnosis biases, which can lead to overestimation of the benefits of screening (4).

In situations in which sufficient randomized clinical trial data are not available, modeling can be used to gain insight into the potential effect of a particular intervention or program on population health outcomes and health care costs. A decision model is used to integrate available data to predict the likely effect of a particular intervention or program in a population, to guide decision making regarding its appropriate use and to highlight issues for further study. Models are of course limited by their dependence on secondary data and multiple assumptions, but they nevertheless have proved useful in a variety of health care settings (5–8). We believe that sufficient clinical data are not currently available to provide a definitive answer regarding whole-body CT screening, and thus we intended our study to be informative rather than conclusive. The purpose of our study, therefore, was to make preliminary estimates of the effectiveness (in life-years) and cost-effectiveness (in costs per life-year) of whole-body CT screening.

MATERIALS AND METHODS

We (M.T.B., G.S.G.) constructed a Monte Carlo (9) decision-analytic model by using software (DATA Pro; TreeAge, Williamstown, Mass) to predict outcomes (in costs and life expectancy) of a large \( n = 500,000 \) cohort of self-referred asymptomatic individuals undergoing onetime whole-body CT screening. Although the protocol for whole-body CT screening is not standardized, a screening examination typically includes unenhanced scanning of the heart, chest, abdomen, and pelvis. We therefore considered the following diseases, which involve eight specific organ systems: ovarian, pancreatic, lung, liver, kidney, and colon cancer; abdominal aortic aneurysm; and coronary artery disease. These conditions are often described on Web sites that advertise whole-body CT screening. We did not include CT colonography, as this is not generally included in whole-body CT studies. The analysis was performed from a quasi-societal perspective; that is, we included costs of disease management, regardless of who incurs them. We did not, however, include costs to the patient, such as time costs (10). We followed the recommendations of the U.S. Panel on Cost-effectiveness in Health and Medicine (11) insofar as was possible, with two exceptions: First, we did not discount future costs or life-years. Because of limitations in data availability, we could use only lifetime costs (from diagnosis to death); it was not possible to determine when costs were incurred, and, thus, accurate discounting of future costs was impossible. Second, we did not account for changes in quality of life.

We performed a base-case analysis in a hypothetical cohort of self-referred 50-year-old men. A base-case analysis is one in which, for each variable included in the model, a value is selected that best represents a real-world scenario (the case). This base value is then used to calculate the principal cost-effectiveness ratio. For sensitivity analyses, we varied the age and sex of the cohort. A schematic representation of the model is presented in the Figure.

Model Structure

Our modeling approach was based on the assumption of a shift in disease stage at presentation as a result of screening, and estimates of survival gains (and costs) were based on the shift toward an earlier stage of disease. It is important to note that without at least this benefit, screening could have no benefit whatsoever and further analysis would be unnecessary. Our modeling approach is similar to that of Marshall et al (12,13), who studied CT screening for lung cancer. Mahadevia et al (14) recently used a similar approach, also to evaluate lung cancer screening. Note, however, that stage distributions, stage shifts, and stage-specific survival and costs were estimated for each of the eight conditions included in our analysis.

We compared two strategies: onetime whole-body CT screening versus no screening (ie, routine care). For whole-body CT screening, the model first assigned each individual a disease or non-disease status for each of the eight conditions on the basis of age and sex.
The model then predicted the results of the CT screening test for each disease on the basis of disease-specific sensitivities and specificities of CT. Thus, in each patient and for each condition, the model classified test results as true-positive, false-positive, true-negative, or false-negative. The model then estimated life expectancy and health care costs based on each patient’s true disease status and test results. Because the model assigned each patient a test result for each of the eight conditions, it allowed patients to have combinations of true and false, positive and negative, test results. Estimated life expectancy was based on the disease condition and test result with the shortest survival, and costs were summed across all disease conditions. For example, a patient might undergo work-up for a false-positive test result for one disease condition, and therapy for another condition on the basis of a true-positive test result. In this case, work-up and treatment costs would be summed, but life expectancy would be assigned on the basis of the combination of test results and true disease status that was associated with the shortest survival.

For routine care, the model assigned patient disease status on the basis of age- and sex-specific disease prevalence in the general population. We chose prevalence as the appropriate measure of disease frequency in this case. Onetime screening would be useful for detecting prevalent cancers, whereas subsequent screening at disease-appropriate screening intervals would allow detection of incident cancers. In the absence of whole-body CT screening, we assumed that prevalence, stage of disease at diagnosis, and health care costs (according to disease and stage) were similar to those in the general population of the same age and sex; since whole-body CT screening is not widely available, we assumed that it would not influence these outcomes in the general population.

We estimated life expectancy and costs with each of these two strategies and calculated incremental cost-effectiveness ratios (ICERs). In calculating the ICERs, we included differences in health care costs between the two strategies in the numerator, and differences in effectiveness in the denominator.

Model Inputs and Assumptions

**Disease prevalence.**—Age- and sex-specific cancer prevalence was estimated on the basis of 1973–1996 data from the Surveillance, Epidemiology, and End Results (SEER) program (15). Data for age- and sex-specific prevalence of abdominal aortic aneurysm and coronary artery disease were obtained from published literature (16,17). We assumed that the presence or absence of each disease was independent of the presence or absence of the other diseases.

**CT performance.**—We (M.T.B., E.W.) derived estimates of disease-specific sensitivity and specificity for unenhanced CT from the published literature (Table 1). Because contrast material–enhanced CT scanning is currently the standard for most organ- or disease-specific CT examinations, we included the most recent publications based on unenhanced CT. If data for unenhanced CT were not available, we (G.S.G.) estimated sensitivity and specificity, taking into account the relatively poorer performance of unenhanced CT.

**Effect of whole-body CT screening on life expectancy.**—In our analysis of the screening strategy, we assumed that cancers that were correctly identified at whole-body CT were detected at earlier stages than if they had been detected incidentally and/or on the basis of the development of symptoms. Without this assumption, screening could have no benefit. Specifically, we assumed that the distribution of stages among cancers detected with whole-body CT would shift approximately 50% toward regional disease from distant disease, and 50% toward local disease from regional disease. A similar approach was used by Mahadevan et al in their analysis of lung cancer screening (14), although we used type-specific stage distributions for each of the six cancers. A weighted average of the stage-specific 5-year survival percentages from the SEER program was calculated by using these adjusted stage distributions. Life expectancy was calculated by using the declining exponential approximation of life expectancy to convert the 5-year survival percentages to life-years (25). To attempt to adjust for lead-time bias, we assumed that 50% of the gain in survival was real and 50% was due to lead-time bias (Table 2). The assumed percentage of apparent survival gain that was due to lead-time bias was varied in subsequent sensitivity analyses.

For patients with whole-body CT–detected abdominal aortic aneurysm, we assumed a life-expectancy gain due to the elimination of rupture risk. Among patients with whole-body CT–detected coronary artery disease (diagnosed on the basis of calcium scoring), we assumed that 20% would initiate statin therapy.
which would cause a 30% reduction in the rate of myocardial infarction in this patient group (26).

We assumed that patients with positive results of scanning (whether true- or false-positive) would undergo additional diagnostic testing, which might include biopsy, to follow up the results of whole-body CT scanning (see the section on Costs for details about the costs of follow-up testing). Patients undergoing biopsy were subject to a mortality risk of 0.001 (1 in 1000) (27). Patients with a false-negative or true-negative result of scanning were assigned an age- and sex-specific life expectancy derived from census data for the general population (1997 U.S. life tables) (28).

Patients with false-negative results of scanning were assigned a life expectancy according to the cancer-specific survival reported by the SEER program investigators or according to the abdominal aortic aneurysm-specific or coronary artery disease-specific survival reported in the literature. We assumed that the clinical course in these patients would be similar to that in patients with no CT screening (Table 2).

Because of insufficient data, no adjustments for health-related quality of life were made in the model. Such adjustments would have required information about quality of life not only for all eight disease conditions but also for all stages of disease for the six cancers. These data were not available in the literature.

Costs

All costs were converted to 2001 U.S. dollars by using the medical component of the consumer price index for that year (29). We (M.T.B.) estimated the cost of a whole-body CT examination (not including CT colonography) from prices advertised on the Internet, which ranged from $800 to $1000. For the purpose of this analysis, we chose a cost of $900.

We derived estimates of health care costs related to disease management from the published literature. When the data were available, we used lifetime costs (ie, costs incurred between diagnosis and death). For ovarian, lung, and colon cancer, we used Medicare reimbursement data according to stage of disease at diagnosis (30,31). We adjusted the costs of disease management for each of two strategies, screening and routine care, by using the corresponding distribution of disease stages. For liver and pancreatic cancer, we used lifetime costs, including costs of follow-up testing and end-of-life care (32,33). For kidney cancer, we used costs of procedures (34). For patients with whole-body CT–detected abdominal aortic aneurysm, we adjusted the costs of routine care by subtracting the costs of emergent aneurysm repair (35,36). For patients with whole-body CT–detected coronary artery disease, we adjusted the costs of routine care for coronary artery disease by subtracting the costs that we predicted would be saved by the prevention of myocardial infarction with the initiation of statin therapy (37) (Table 3).

We did not discount future costs in the model because our use of lifetime aggregate costs of disease management limited our ability to determine when costs were incurred. Furthermore, since discounting must be applied equally to both costs and effects, we also did not discount life-years.

Costs of follow-up testing for patients with false-positive scans were derived from Medicare reimbursement rates to physicians and hospitals on the basis of Common Procedural Terminology codes (38) (Table 3). We assumed that all biopsies were performed with CT guidance, with the exception of colon cancer biopsy, which we assumed was performed by using colonoscopy. We assumed a 1% risk of complication (eg, perforation, collapse of lung, or severe bleeding) for patients who underwent biopsy, and an estimated cost of $5000 for a 3-day hospital stay associated with complication of biopsy (39).

Sensitivity Analyses

By means of sensitivity analyses, we (M.T.B.) examined how variation in the values of model parameters affected the cost-effectiveness results. We varied the disease prevalence in the screening population, the proportion of life-expectancy gain attributable to lead-time bias, and the sensitivity and specificity of whole-body CT. We also varied the costs of follow-up for false-positive screening results, as well as those of whole-body CT and disease management. Finally, we varied the rate at which biopsies were performed and the risk of complications. We performed all sensitivity analyses for both men and women aged 45–55 years.

RESULTS

Base-Case Analysis

Using our base-case model estimates, screening with whole-body CT in 50-
Sensitivity Analyses

The results of our analyses were sensitive to disease prevalence, the effect of screening on stage of diagnosis, the specificity of whole-body CT, the cost of follow-up for false-positive findings, and the cost of the whole-body CT examination. For example, when the prevalence of all diseases considered was reduced by 50%, the ICER of whole-body CT in 50-year-old men increased to $301 000. In the same cohort, a 50% increase in disease prevalence reduced the ICER to $97 000. A 50% increase in the proportion of life-expectancy gain attributable to lead-time bias resulted in an ICER of $204 000, and, conversely, a 50% decrease in the proportion of life-expectancy gain attributable to lead-time bias resulted in an ICER of $153 000 (Table 4).

A 5% increase in the specificity of whole-body CT for the detection of all eight disease conditions resulted in reduction of the ICER to $125 000, largely because of the reduction in the number of false-positive test results. The results were less sensitive to the sensitivity of whole-body CT screening for disease detection. For example, a 5% increase in the sensitivity of whole-body CT resulted in a decrease in the ICER to $148 000.

Decreasing the costs of follow-up testing by 50% reduced the ICER to $96 000, and increasing the costs of follow-up testing by 50% increased the ICER to $203 000 among 50-year-old men. The effect of whole-body CT costs was less pronounced; reducing the cost by 50% resulted in a reduction of the ICER to $124 000 among 50-year-old men.

Whole-body CT screening was consistently less cost-effective for women than for men, largely because of the lower overall prevalence of these diseases in women. For example, whole-body CT screening cost $151 000 per life-year gained for 50-year-old men but cost $170 000 per life-year gained for 50-year-old women. This trend held true across all ages considered. Among 45-year-olds,
whole-body CT cost $194 000 per life-year gained for men and $267 000 per life-year gained for women (Table 4).

Factors that had little effect on the results included the costs of caring for patients with disease, the rate at which biopsies were performed for false-positive test results, and the risk of complication of biopsy.

**DISCUSSION**

The results of this analysis suggest that in a population of asymptomatic 50-year-old men, onetime screening with whole-body CT would cost an additional $151 000 per life-year saved, relative to survival with no screening. Our findings suggest that whole-body CT is more expensive, in costs per life-year, than the majority of health care interventions currently funded in the United States (40). For example, center hemodialysis, which Medicare is required by law to fund, has been estimated to cost between $55 000 and $80 000 per life-year gained (41).

Our results were most sensitive to the overall prevalence of disease in the screening population, the effect of screening on stage of diagnosis, the specificity of whole-body CT, the cost of additional testing for false-positive whole-body CT results, and the cost of the whole-body CT examination. The results were relatively insensitive to the risk of complications of biopsy for follow-up tests, the rate at which biopsy was performed in patients with false-positive test results, and the costs of disease management. Our estimate of a high ICER suggests that whole-body CT screening is not appropriate in a population with an average risk (ie, low prevalence) of disease. Whole-body CT appears more cost-effective as overall disease prevalence increases; however, our comparison of screening with no screening may be inappropriate in a higher-risk group of individuals, in whom a comparison of whole-body CT with another form of screening or surveillance would be more relevant.

The model’s sensitivity to whole-body CT specificity and to the cost of additional follow-up in patients with false-positive examinations is not particularly surprising. With an increase in the number of false-positive results (decrease in specificity) or in the cost of patient work-up for these results, the ICER of screening should increase. Our base-case analysis suggests that only 17% of the cost associated with widespread implementation of whole-body CT screening would relate to the cost of scanning, and 32% of the total cost would be attributable to work-up for false-positive test results. From a policy perspective, this finding makes clear that the financial effect of whole-body CT extends far beyond the test itself, and therefore a true estimate of total costs of whole-body CT screening must include downstream costs.

Our methods and results are similar to those of a recent cost-effectiveness analysis of CT screening for lung cancer by Mahadevia et al (14). In particular, our study and that of Mahadevia et al were based on a stage-shift assumption; both used similar estimates for the magnitude of the stage shift and both adjusted for lead-time bias. Mahadevia and colleagues found that screening for lung cancer in smokers cost $116 300 per quality-adjusted life-year, compared with no screening in the same population. Our higher (ie, less favorable) estimate of the cost per life-year gained is understandable because of the lower overall prevalence of disease in the likely whole-body CT screening population (lung cancer screening is generally performed in a population with higher risk), the expected increases in false-positive results and associated work-up costs as a result of the evaluation of a larger number of conditions, and the relatively poorer performance of CT for the detection of several of the cancers sought in a whole-body CT examination. It is important to note that we allowed false-positive results for each of the eight conditions included. If screening were implemented, additional false-positive results (and perhaps true-positive results, as well) might be seen also for other conditions. Because other conditions are less common and because the false-positive and true-positive results likely would have offsetting effects, we do not think our limit of eight conditions substantially affected the results.

The principal limitations of our study relate to the uncertainty inherent in some of our model parameters and to the simplifying assumptions necessary to develop a tractable model. An important area of uncertainty is the survival gain associated with earlier detection of disease with screening. Specifically, at least part of the increase in survival estimated on the basis of cancer stage shift would be artificial because of lead-time bias. Our approach to estimating the shift in type-specific stages, likely gains in survival, and the effects of lead-time bias is similar to those reported for previous studies. In the absence of clinical trial data, however, it is difficult to predict the magnitude of these effects. There is further uncertainty in model parameters such as disease prevalence, CT performance, effect of screening on life expectancy, and test and treatment costs. We attempted to provide reasonable estimates for each of the specific model parameters and probability distributions and to justify these with references to the published literature and expert opinion. Furthermore, we attempted to err on the side of favoring whole-body CT—not because we are in favor of it, but to avoid biasing the results against it and to establish conditions under which it might or might not be cost-effective.

The model developed for these analyses is a necessary simplification of reality, as are all cost-effectiveness models. It is not possible to model all diseases that might be detected with whole-body CT screening or all possible combinations of follow-up tests, results, and outcomes. We chose to limit our analysis to eight specific diseases and representative follow-up strategies for each test result and disease state combination. In addition, we did not account for potential deleterious effects of radiation exposure from CT. These may be particularly relevant in patients who undergo multiple follow-up scans for a false-positive result or who choose to undergo repeat whole-body CT. We assumed that all biopsies would be performed with CT guidance. We believed that this simplifying assumption was justified, given that biopsy was performed in only a small percentage of the group of patients with positive results (10%) and that to include biopsies performed with another imaging modality (eg, US) would only affect costs. In sensitivity analyses, we varied the rate at which biopsy was performed and the risk of complications due to biopsy, and we found that neither parameter influenced the cost-effectiveness results. In the model, we did not discount future costs or life-years. We also assumed independence between all eight conditions. While it may be that there are some relationships between some of the conditions, we did not include them in our analysis because of the lack of sufficient data on the subject and the very low overall prevalence of these conditions in an asymptomatic self-referred population.

We did not evaluate the effect of race, income, or education on screening effectiveness or cost-effectiveness. These characteristics are known to correlate with disease prevalence and disease-specific
survival, and they should be included in future analyses. For example, it may be that individuals who undergo whole-body CT screening have higher socioeconomic status or are more health conscious (the “worried well” effect) than the general population. In this case, their lower disease prevalence or higher disease-specific survival in the absence of screening would reduce the incremental benefits and thereby increase the ICER of whole-body CT screening.

We evaluated the cost-effectiveness of onetime whole-body CT screening. We believe that the onetime screening examination is most representative of the whole-body CT market at this time; it is possible, however, that whole-body CT in future might be used in a manner more typical of other screening tests, or, in other words, might be performed at regular intervals. We believe that a onetime CT examination would be more likely to be cost-effective than would multiple examinations, as initial scanning would enable detection of prevalent cases of disease, which are presumably greater in number than are incident cases in the same population. In addition, the costs associated with false-positive test results and risks from unnecessary radiation exposure would only be magnified with the addition of subsequent CT examinations.

With respect to the cost of a whole-body CT examination, we used a base-case estimate of $900, which is based on prices advertised on the Internet. It is possible that the cost per examination is less than $900 or that the price will decrease over time. Screening tests are generally less expensive than are diagnostic examinations of the same body part (eg, screening vs diagnostic mammography), and whole-body CT screening examinations are currently priced at a level roughly comparable with diagnostic CT examinations. Even if the cost of whole-body CT were to decrease by 50%, however, the ICER for CT screening would not be less than $100 000 per life-year except in the relatively higher-risk group of 55-year-old men.

Finally, we did not attempt to quantify or include in our analysis any positive or negative psychological effects of false-positive or true-positive results of CT screening. While we recognize the possibility that such effects might be seen, we know of no studies that have measured their impact on quality of life by using preference-based methods. In addition, the outcomes of our analyses were costs and life-years, rather than quality-adjusted life-years, which is the category in which psychological effects would be seen. Even if we had attempted to incorporate these effects into our analysis, the benefit to quality of life with the peace of mind that comes from knowing one is cancer free may potentially be offset by the detriment caused by anxiety due to a false-positive test result.

In conclusion, ultimately policy with regard to the implementation and management of whole-body CT screening must be based on judgments concerning the likely benefits of screening to the population and the costs of achieving these benefits. We also attempted to define conditions in which whole-body CT screening might or might not be cost-effective when compared with no screening. Our analysis was performed to estimate the benefits, the costs of achieving those benefits, and the incremental cost-effectiveness of whole-body CT screening relative to routine care. Our results, although preliminary, suggest that whole-body CT screening would result in substantial costs with minimal population-wide gains in life expectancy. These findings are due principally to the low overall prevalence of disease in the population likely to undergo screening and the cost of follow-up tests for patients who have false-positive screening results. Furthermore, the majority of the costs attributable to whole-body CT screening would accrue to the health care system rather than to the individual undergoing screening (ie, insurance would cover most follow-up tests and treatments). In sum, we believe that this analysis has raised questions about the potential cost-effectiveness of whole-body CT screening in a self-referred population relative to other currently funded health care programs and that further serious consideration should be given to the costs and benefits of this technology before it is more widely used.

References
5. Prosser LA, Stinnett AA, Goldman PA, et al. Cost-effectiveness of cholesterol-lowering therapies according to selected pa-