Economic issues in the use of office spirometry for lung health assessment

M Murray Krahn MD MSc¹, Kenneth R Chapman MD MSc²

The National Lung Health Education Program (United States) has recently recommended using office spirometry to screen for subclinical lung disease in adult smokers. No published studies evaluate the economic consequences of this recommendation. This review article outlines the issues that must be considered when evaluating the costs and health benefits of office spirometry. Much of the available data on the effectiveness of screening is from studies that included smoking cessation interventions, making it difficult to determine the effects of screening alone. The sensitivity and specificity of screening spirometry are not known, but may not be important in the economic model, because even false positive test results are beneficial if they lead to smoking cessation. Costs to be considered include those of spirometry itself, of implementing and maintaining screening and smoking cessation programs, and of their consequences, ie, reduced morbidity (lower short term health care costs) and mortality (perhaps higher long term health care costs). Despite these unique challenges, data are available to perform economic analyses regarding screening spirometry. Such analyses should play a role in future clinical policy making. Even modest quit rates attributable to screening spirometry may result in highly favourable cost effectiveness ratios.

Key Words: Cost effectiveness; Cost of illness; Economics; Pulmonary function tests; Screening; Spirometry

Using spirometry to screen smokers for evidence of subclinical obstructive airways disease is intuitively appealing. Chronic obstructive pulmonary disease (COPD) incurs a large clinical and economic burden of illness. It is now the fourth leading cause of death in the United States, surpassed only by heart attacks, cancer and stroke. A similar number of individuals die from COPD per year (n=100,000) (1) as from lung cancer (n=119,000) and smoking-related cardiovascular disease (n=180,000) (2). Smoking-related illness accounts for 6% of all direct health care expenditures in the United States; respiratory disease represents approximately one-third of this total (1,3).

Furthermore, we know that spirometric testing is able to select those smokers who are at highest risk for developing COPD (4-6). We also know that smoking cessation is an effective means of preventing progression to clinical COPD. There is both observational (5,7) and experimental (6) evidence suggesting that the rate of forced expiratory volume in 1s (FEV₁) decline and therefore the risk of progression to COPD is decreased in smokers who quit. The National Lung Health Education Program, on the basis of existing evidence, has launched an initiative to promote the use of office spirometry for screening in smokers older than the age of 45 years (8). The British Thoracic Society guidelines also suggest that smokers, particularly those with occupational risk, be screened (9). However, other specialty group guidelines (10,11) and evidence-based guidelines put forward by the Canadian Task Force on the Periodic Health Examination (12), the American College of Physicians (13), the Agency for Health Care Policy and Research (14) and the American Thoracic Society (15) do not recommend that smokers be screened. Why the disagreement?

Department of Medicine, ²Health Policy, Management and Evaluation, University of Toronto, Toronto, Ontario
Correspondence: Dr M Krahn, The Toronto General Hospital, ENG-248, 200 Elizabeth Street, Toronto, Ontario M5G 2C4.
Telephone 416-340-4155, fax 416-595-5826, e-mail murray.krahn@uhn.on.ca

Can Respir J Vol 10 No 6 September 2003
One reason is cost. There are approximately 60,000,000 smokers in the United States and Canada. At $20 per test, screening would incur direct costs of over $1 billion per episode. Follow-up visits, repeat office spirometry, full pulmonary function tests, lung imaging, drug prescriptions and smoking cessation interventions would follow initial office spirometry in many patients. Given current constraints on health expenditures, careful examination of the evidence for the effectiveness and cost effectiveness of screening spirometry is required before promulgation of clinical policies whose impact on aggregate health expenditures is so large. Highlighting the importance of this problem, a recent workshop sponsored by the National Heart, Lung, and Blood Institute in the United States indicated that the economics of smoking cessation, and in particular the cost effectiveness of screening for COPD, were key priorities in developing national research and prevention strategies for COPD (16).

How would one go about evaluating the economic attractiveness or ‘cost effectiveness’ of screening spirometry? Standard texts on economic evaluation in health care are available (17,18). However, in this review we consider some of the unique issues in the design of an economic study of spirometric screening and then see how well the existing literature has addressed them.

COMPARISONS

Economic evaluations are usually comparative, and full evaluations are always so. Though a simple description of costs and consequences of a single program may be a useful contribution to knowledge, it is insufficient to allow a judgment about its economic attractiveness. An intervention is always more or less economically attractive in relation to some other program, even if the other program is ‘no intervention’. Thus, screening spirometry will be more or less attractive in comparison with something else: either ‘no intervention’, ‘usual care’ or a smoking cessation program that does not incorporate screening spirometry.

Selection of either the intervention or the comparator(s) is not entirely straightforward. The intervention may be simple office spirometry or full pulmonary function testing. The intervention may be delivered to smokers who are presenting for periodic health examination, for respiratory complaints or for unrelated health problems. This type of ‘screening’ is often referred to as ‘case finding’ (15). The term ‘screening’ is usually reserved for population screening programs in which individuals are screened at a work site or other public site (eg, shopping mall), or are recruited to a central location through public advertisements. Population screening and case finding may differ in cost, efficacy, and evidentiary and ethical standards that must be met before implementation (15).

Attempts to define comparator and intervention are often complicated by the fact that individuals with abnormal spirometry results frequently receive additional interventions, whose effects are difficult to disentangle from the tests themselves. In the Lung Health Study (6), for example, all study participants received screening spirometry and a physician-delivered, patient-specific message that the individual was at high risk for COPD. In addition, some patients received a multifaceted behavioural and pharmacological intervention targeted at smoking behaviour. The difference in quit rates between the two groups represented, therefore, is the aggregate effect of multiple factors.

Because most of the debate concerning the economics of screening spirometry has revolved around the screen itself, the incremental costs and health benefits of screening alone should be evaluated. Thus, both ‘screening’ and ‘comparator’ groups should receive the same behavioural and/or pharmacological antismoking interventions. Expert panels have issued guidelines suggesting that all smokers, not only those at high risk, should be targeted for antismoking interventions (14). Therefore, the value of screening spirometry should be judged by its ability to produce an increment above the quit rate observed in standard smoking cessation programs.

Alternatively, one may model a strategy that more closely mirrors clinical practice and that takes compliance with published guidelines into account. In clinical practice settings, smoking cessation efforts in unselected smokers may be modest. Abnormal spirometry tests may trigger more intensive, multifaceted and possibly effective smoking cessation efforts on the part of the physician and better compliance on the part of the patient. Whatever approach is chosen, the effects of the screen itself and the ancillary interventions that follow the screen must be kept conceptually distinct.

The population to whom the screening intervention is delivered must be considered. An unselected population including all smokers represents one option. Another option is to screen only those with cardiovascular risk factors such as diabetes, hypercholesterolemia or strong family history. Although this latter approach may seem a bit counterintuitive (ie, using cardiovascular risk factors to select patients for lung health screening), the health benefits per person who stops smoking will likely be proportionately greater, because the benefits of smoking cessation are primarily cardiovascular.

STUDY ARCHITECTURE: MODELLING AND CLINICAL TRIALS

There are two fundamentally different approaches to gathering and combining the cost and health benefit data in an economic evaluation. The first approach is to gather economic data simultaneously with clinical data in a randomized, controlled trial designed to evaluate the effectiveness of screening. In this ‘piggyback’ approach (19-21), the resource implications of gaining extra health benefits are computed directly from trial data at completion. The benefit of this approach is high internal validity of efficacy and cost estimates. A major drawback is the limited generalizability of trial results, particularly when practice patterns and costs are unrepresentative of community practice.

The other approach is to compare screening alternatives using decision analytical or simulation models (22,23). In this approach, cost data and health data are gathered from disparate (usually secondary) sources and incorporated into a common model. Compared with a clinical trial, modelling allows the evaluation of many more alternatives and longer time horizons, but the comparisons between programs may be subject to greater bias.

Future economic studies of screening spirometry may adopt either approach. Because the Lung Health Study has provided
a new legitimacy to the idea of screening smokers, a retrospective analysis using data from this study seems logical. However, this trial considered a single strategy that may not be widely generalizable, ie, a very intensive, multifaceted, behavioural and pharmacological smoking cessation strategy delivered over five years. It does not directly address the value of screening, but rather the effect of a complex intervention in a high risk group identified by screening. To determine the value of screening alone, the costs and health effects of a cohort of nonscreened smokers, as well as the cohort of screened smokers who were excluded from the trial, would also need to be considered. In addition, trial participants were highly unrepresentative of the general population. They volunteered to participate not only in a clinical trial, but also in an intensive smoking cessation program. Finally, additional modelling work would be required to fully assess the benefits of smoking cessation achieved in the Lung Health Study, because these extend over decades.

An alternative approach uses decision analytical or simulation models to consider the effects of screening spirometry in conjunction with a variety of smoking cessation interventions. A greater number of program alternatives, including ongoing or repeat screening and smoking cessation interventions that more closely mirror clinical practice, could be considered in an expeditious manner and at a reasonable cost.

**TEST PERFORMANCE CHARACTERISTICS**

Archie Cochrane, a British epidemiologist whose name graces the international research effort known as the Cochrane Collaboration, suggested that a good screening test should be simple, acceptable, reliable, valid, sensitive, specific and predictable (24). Office spirometry appears to meet most of these criteria. What about sensitivity and specificity?

First, any attempt to evaluate the operating characteristics of screening spirometry in the context of a full economic evaluation involves assumptions about the natural history of obstructive airways disease. We simply do not have data that directly describe the sensitivity and specificity of the test, ie, the probability of developing clinical obstructive airways disease given a particular FEV₁ or FEV₁/forced vital capacity (FVC) result some decades earlier. The outcome measure for most prognostic studies in patients with early air flow obstruction, including the Lung Health Study, is the rate of decline in lung function, eg, FEV₁ or FEV₁/FVC ratio (4,5,7,25,26). These represent secondary outcome measures, and the rates of occurrence of primary end points such as mortality and clinical obstructive airways disease can be predicted only using models of the natural history of disease.

One very simple approach to thinking about this problem would be to derive sensitivity and specificity values from a plot of FEV₁/FVC versus rate of FEV₁ loss. A positive test could be considered to be a FEV₁/FVC ratio less than 70% to 75%. Those who are ‘disease positive’ may be those with a rate of FEV₁ loss greater than 50mL/year or 75 mL/year. The sensitivity of the test could be described as the proportion of those with abnormal spirometry whose rate of FEV₁ loss exceeds a preset threshold. Sensitivity analysis could evaluate the optimal threshold for screening spirometry.

This approach amounts to a restatement of the ‘horse racing hypothesis’, bruited about in the 1980s, which stated that those with abnormal spirometric results lose lung function more rapidly and are therefore at higher risk of obstructive airways disease than those with normal spirometry. Some evidence for this hypothesis was found in male, but not in female, smokers (4).

A more sophisticated approach would dispense with sensitivity and specificity entirely. The population of screeners and nonscreeners could be divided into multiple strata, characterized by a series of thresholds (of FEV₁, for example). The annual probability of changing strata for smokers, nonsmokers and ex-smokers would be modelled as a function of the mean rate of FEV₁ loss, given smoking status and initial spirometric results. The rates at which primary endpoints (COPD, COPD-related death, non-COPD death) occur would be linked to a given threshold of pulmonary reserve (eg, FEV₁). This approach relies on the conceptual model of the natural history of COPD first proposed by Fletcher and Peto (5). It implicitly incorporates the value of screening spirometry in two ways: spirometry would predict the rate of lung function loss and the remaining reserve at the time of screening. Thus, sensitivity and specificity are implicitly present and could be derived post hoc, but are not explicitly present as a variable in the model.

One insight about sensitivity and specificity that emerged from our early efforts to model the effects of screening spirometry, but has also been commented on by others (27,28), is that sensitivity and specificity do not matter or, at least, they do not matter much. The consequences of any positive test, including a false positive test, are beneficial, because even smokers who are not truly at risk of COPD derive a large health benefit from quitting. These benefits include reduction in coronary heart disease and lung cancer incidence. Thus, sensitivity analyses in future economic evaluations will probably show that analytical results are not particularly sensitive to the screening threshold. They will likely also show that the optimal approach to setting thresholds for spirometric test abnormality is to call every single test a smoker positive! This is a formal restatement of the ‘horse racing hypothesis’, bruited about in the 1980s, which stated that those with abnormal spirometric results lose lung function more rapidly and are therefore at higher risk of obstructive airways disease than those with normal spirometry. Some evidence for this hypothesis was found in male, but not in female, smokers (4).

EFFECTIVENESS

Future analyses, however, are quite likely to be sensitive to the effectiveness of screening, expressed as the incremental quit rate in those who are screened relative to those who are not. What do we know about the effectiveness of screening spirometry? Not much – few randomized studies directly address the effect of screening spirometry on quit rates. Risser and Belcher (29) compared spirometry, carbon monoxide and symptom-related counselling with standard smoking cessation advice in 90 patients in a randomized trial at an outpatient general medicine clinic. Although multiple interventions were
delivered, only one group received spirometry testing and results. At 12 months, the carbon monoxide-validated smoking cessation rate was 23% higher in the screened group than in the controls (33% versus 10%, P=0.03).

Segnan et al (30) randomized 923 smokers in primary care settings in Italy to counselling with spirometry, counselling alone and brief physician advice. The quit rates (6.5%, 5.5% and 4.5%, respectively) were not statistically significantly different. A population-based study in Norway (31) identified 2610 young men with a history of asbestos exposure and low FEV₁ values. One-half of them were mailed a personal letter from their physician explaining their risk and advising them to stop smoking. In addition, the intervention group received a smoking cessation pamphlet. Crude quit rates were 5.6% and 3.5% in the intervention and nonintervention arms, respectively, a difference that remained statistically significant after adjustment for smoking intensity and asbestos exposure.

Nonrandomized studies show similarly heterogeneous results. Hepper et al (32) reported the results of a community lung screening program organized by the Minnesota Lung Association. The intervention consisted of advertising in selected communities, screening in a mobile van, advice to patients with abnormal results to see their physicians and a series of educational seminars for community physicians. The excess quit rate at two to three years in smokers with abnormal lung function was approximately 10% (32).

In the Lung Health Study, approximately 22% of the smoking intervention groups were sustained quitters at five years, compared with 5% in the usual care group (6). The excess quit rate is comparable with the 14% excess quit rate in smoking programs with more than two interventions reported in a meta-analysis of all smoking cessation interventions (33). However, the difference in quit rates in this study is due to the smoking intervention, not spirometry, because all patients received spirometry and were told of their abnormal results. The effect of spirometry alone appears less impressive when the 5% quit rate in the control group (who received spirometry) is compared with the expected 1% to 3% per year ‘natural’ quit rate among all smokers.

Given these scanty and heterogeneous data, it may be difficult to arrive at precise estimates of the incremental effectiveness of spirometry alone in future economic studies. Indeed, one recent evidence-based review suggested that there is no conclusive evidence that spirometry adds to the efficacy of standard smoking cessation advice.

We do not offer a summary judgment here about the state of the evidence. However, we make the following observations. First, screening spirometry must have some incremental benefit in improving quit rates for it to be justified as a rational health policy. Simply identifying smokers at high risk is a waste of resources unless it leads to improvement in health status. Such a strategy cannot be cost effective; cost effectiveness depends on effectiveness.

Second, even very modest quit rates, if they are in fact the consequence of spirometry, may be sufficient to render screening spirometry ‘cost effective’. Published cost effectiveness analyses have shown that smoking cessation interventions with incremental quit rates of 3% to 6% are quite economically attractive despite their unimpressive effectiveness at first blush (34-36). The reason why even modest quit rates are sufficient has to do with the very large health benefits associated with smoking. A small proportion of health benefit is attributable to reduction in obstructive airways disease. A much larger proportion of health benefit is attributable to reductions in coronary heart disease and lung cancer incidence rates.

Finally, there is the issue of emphasis. The recent special report of the National Lung Health Education Program (8) devotes pages of text to the issues of precision, reliability and calibration of office spirometers. The issue of what to do with abnormal results is not mentioned. For office spirometry to be effective and cost effective, it must be directly linked to specific, effective, preferably multifaceted smoking cessation interventions targeted to those at highest risk. Otherwise, purchasing, calibrating, maintaining and using office spirometers for screening is almost certainly a waste of time and resources.

In addition, a comprehensive economic evaluation must consider not only the intervention itself, but the effects on the case mix receiving the intervention. For example, referral of high risk smokers identified in a screening program to smoking cessation interventions may improve their aggregate health output by increasing the total number of smokers who quit, causing more high risk smokers to quit or both.

COSTS

Table 1 is a partial list of costs that should be included in a full economic evaluation of screening spirometry. The initial costs associated with screening include not only the costs of the intervention itself, but also the costs associated with program infrastructure and advertising, and induced costs related to more intensive medical care in screenees. These costs are very frequently excluded in analyses of screening programs (37).

Little has been published about the overall costs of screening spirometry. Owens (38) estimated the costs of recruitment in the Lung Health Study (ie, screening and other trial costs excluded) to be US$14 per abnormal result in work sites and public sites. Costs for radio, direct mail and newspaper advertising were approximately US$150 per abnormal result, and costs for advertising on public transit increased to US$778 per abnormal result (38).

Better information is available about downstream costs. The economic consequences of smoking have been described in numerous reviews (2,39-41). Most of this information, however, is not directly applicable to a screening population, because the effects of screening are largely seen only in high risk smokers, ie, those with abnormal test results. Short term health expenditures are likely be higher than among all smokers and long term expenditures probably lower due to higher mortality rates. The net effect is difficult to predict in the absence of modelling.

Some limited cost data on caring for COPD patients are available (42-48), but future studies may require additional empirical costing work to accurately predict the long term economic effects of spirometric screening.

HEALTH EFFECTS

Table 2 is a partial list of health effects induced and averted by spirometric screening. The short term psychological effects of screening in other clinical contexts are real but dif-
and have been assessed with generic utility instruments (35,57). An alternate approach excludes all quality of life considerations and focuses only on the mortality benefits of screening in a cost-benefit or cost effectiveness (cost per year of life saved) analysis. This much simpler but less comprehensive approach is often taken in screening studies in other contexts (63-66) and in economic analyses of smoking cessation strategies (34,35).

### ECONOMIC EVALUATION OF SPIROMETRIC SCREENING

In the only published economic evaluation of lung function testing, Loss et al (67) screened 73 young smokers, 75% of whom had abnormal pulmonary function test results. At six months, 7% of those with abnormal tests had quit smoking. The cost per incremental quitter was estimated to be US$1392. No control group was assessed, and the long term health and economic effects of smoking cessation were not considered. This result is therefore of marginal value in policy formulation.

### CONCLUSIONS

Although spirometry is relatively inexpensive, the aggregate economic and health effects of testing all smokers are large. Economic studies should play a role in future clinical policy formulation regarding the role of screening spirometry. Although economic studies in this area face some unique challenges, good quality cost and outcome data are available for many components of the problem. Several good quality economic analyses of smoking cessation programs (34-36) could provide a template to researchers beginning to design economic analyses of screening spirometry. As these studies suggest, even modest incremental quit rates could show screening spirometry to be quite cost effective.

### ACKNOWLEDGEMENTS:
Support for this project was provided by an unrestricted grant from Boehringer-Ingelheim.
REFERENCES


Submit your manuscripts at http://www.hindawi.com