Occasional essay

Screening for lung cancer: Can it be cost-effective?

Olli S. Miettinen

Abstract

LAST YEAR, THE LANCET REPORTED ON A STUDY CONCERNING A PARTICULAR, AVANT-GARDE REGIMEN OF CT-BASED SCREENING FOR LUNG CANCER, SHOWING ITS GREAT SUPERIORITY RELATIVE TO THE CORRESPONDING REGIMEN BASED ON TRADITIONAL RADIOGRAPHY (LANCET 1999;354:99-105). THAT REPORT WAS MET WITH GREAT INTEREST IN THE MEDIA, WORLDWIDE. IT THEREBY ALSO LED TO SUBSTANTIAL PUBLIC DEMAND FOR THE STATE-OF-THE-ART SCREENING, IN THE UNITED STATES ESPECIALLY. I HERE ARGUE THAT, DESPITE THE PREVAILING OFFICIAL RECOMMENDATIONS AGAINST ANY AND ALL SCREENING FOR LUNG CANCER IN THE UNITED STATES AND CANADA, IT ACTUALLY ALREADY IS KNOWABLE THAT MODERN SCREENING, WITH SUITABLE SPECIFICATIONS OF BOTH THE SCREENING ITSELF AND ITS RECIPIENT, NOT ONLY IS EFFECTIVE BUT CAN BE EFFECTIVE ENOUGH TO AMPHILY JUSTIFY ITS COST. IT THUS IS TIME FOR AUTHORITIES TO FORMULATE, EVER MORE INCLUSIVELY, THOSE COST-JUSTIFYING SPECIFICATIONS — AND TO PROMOTE RESEARCH PROVIDING FOR FURTHER EXPANSIONS OF AND INNOVATIONS IN THEM. AMERICAN AUTHORITIES, HOWEVER, HAVE NOT REACTED THIS WAY TO THE NEW SITUATION AND, I DARE SAY, THEY HAVE BEEN TENACIOUSLY IRRATIONAL (AND THEREBY IRRESPONSIBLE) IN THEIR UNDERLYING IDEAS ABOUT THE NATURE OF THE PROPER KNOWLEDGE-BASE FOR SCREENING PRACTICE AND OF THE RESEARCH SERVING TO ADVANCE THIS. IT REMAINS TO BE SEEN HOW TIMELY, AND HOW COMPELLINGLY RATIONAL, THE CANADIAN OFFICIAL REACTIONS WILL BE; AND THIS TOO MATTERS GREATLY, AS COUNTLESS LIVES HANG IN THE BALANCE, WITHIN CANADA AND, ESPECIALLY, IN THE WORLD AT LARGE. EX CANADA LUX?

A particular regimen of screening, when applied to people of a particular kind at a given cost, may or may not be cost-effective. If it is the belief of experts that it is, then that regimen should be recommended for use in those people and such practice should be covered by health insurance. Analogously, if the belief is that it is not cost-effective, then it should be recommended against and should not be covered by insurance. And if experts are in doubt or in disagreement, there should be no recommendations for or against, no guidelines in respect to that particular combination of types of screening and screenee — and no presumption of knowledge-based policy in respect to insurance coverage.

As for lung cancer, eminent agencies and organizations in the United States and Canada have recommended against any type of screening on any type of indication, and this policy continues to be in place; but meanwhile, recommendation for lung cancer screening has been adopted as a public-health policy in Japan. Accepting cost-effectiveness as the basis of rational policy and presuming the attainability of cost-effective screening to be essentially the same in the United States or Canada as in Japan, it must be concluded that at least one of these conflicting policies is irrational — and that both of them are if, even in respect to the most promising combination of types of screening and screenee, it remains unknown whether acceptable cost-effectiveness is attainable in efficient practice.

The question of whether screening for lung cancer, with suitable specifications, can now be taken to be cost-effective has suddenly become one of keen interest, following the report that my colleagues and I published of our baseline experience with CT screening in the Early Lung Cancer Action Project (ELCAP). That report principally showed that a particular avant-garde regimen of CT-based screening is, relative to its counterpart based on traditional radiography ("x-ray"), greatly superior in "detecting" (i.e., prompting the diagnosis of) asymptomatic lung cancer; and further, that false-positive results of the CT screening test, under suitable guidelines, do not constitute a serious problem even in baseline screening (in which they naturally are much more common than in repeat screening, the latter focussing on new findings). The editorial commentary attached to the report had a telling title: "Screening for lung cancer: time to think positive." The media took great interest in this, worldwide; and in the United States in particular, considerable public demand for the screening immediately arose. At the same time, though, the US National Cancer Institute counselled caution and proceeded to develop plans for a long-term randomized controlled trial (RCT) on the lifesaving effectiveness of "screening" (i.e., of early intervention provided for by early, screen-prompted diagnosis in place of symptom-prompted diagnosis).

The issues surrounding screening for lung cancer are not only suddenly topical; they are also exceptionally big, as this cancer in North America now causes more deaths than colon, breast, uterine and prostate cancers combined.

While the US National Cancer Institute apparently will, for quite a while, be studying whether screening for lung cancer is at all useful in the prevention of deaths from this disease, near-universally fatal in the absence of screening, perhaps the authorities here in Canada will, without any undue delay, concern themselves with the question of: with...
How to think about the cost-effectiveness?

Any given regimen of screening for lung cancer needs to be defined in respect to the following elements: the screening test; positive results of the test; the diagnostic algorithms that are activated, respectively, by various positive results of the screening test; and the frequency with which such unit “screening” (i.e., the test together with the result-contingent definitive diagnostics) is repeated. The screening — pursuit of early diagnosis, that is — at baseline is prone to differ from the repeat screenings; and if it does, it is the nature of the repeat screenings that in all essence alone has bearing on the regimen’s cost-effectiveness. After all, application of the screening test in any given individual is prone to be repetitive and is, thus, typically of the repeat-screening type.

The unit screening’s successive repetitions in a given individual do not have identical effectiveness. The probability (prevalence) of asymptomatic-but-detectable cancer changes (increases) over increasing age, and also the detection’s utility in extending life expectancy changes (decreases) over increasing age. For this reason already, the regimen’s cost-effectiveness in the context of a particular indication (profile implying probability of presence of latent lung cancer jointly with age per se, mainly) is best thought of in reference to a single repeat screening. This focus is essential for optimization of the regimen’s application to any given individual — mainly a matter of defining the ages at which to begin and end the screenings. Moreover, it provides for greatly enhanced simplicity in dealing with cost-effectiveness.

In the cost (monetary) of a single repeat screening, the main element is the cost of the screening test itself. When false-positive results of the test (in its repeat application) are not very rare, one relevant added element is the typical cost of the diagnostic work-up brought about by a false-positive result of the test, multiplied by the probability of such a result. It is this ruling-out work-up in the absence of diagnosable cancer that, apart from the screening test itself, mainly is peculiar to screening. For, when diagnosable cancer is present, the screen-prompted diagnosis merely substitutes for what otherwise would become a later, symptom-prompted diagnosis. There is, however, a positive difference of cost between screen-prompted early diagnosis and symptom-prompted later diagnosis of the rule-in type, and this difference multiplied by the probability of screen-prompted diagnosis (rule-in) is yet another element in the cost (expected) of a screening (single, repeat). And finally, the costs of intervention are different between screen- and symptom-diagnosed cases — distinctly lower with screen-diagnosed cases, in which resection usually is all that is required — and this difference (negative), again multiplied by the probability of screen-prompted diagnosis, enters into the cost.

The effectiveness associated with this cost (of a single repeat screening) — the gain that might justify the cost — is, as with the effectiveness of whatever purposive action, the extent to which it produces its intended consequence. Given the nature of the screening regimen, together with the profile of the screenee at the time of this screening (notably smoking history and age), there is a corresponding probability for diagnosis (rule-in) of lung cancer resulting from this particular instance of the regimen’s application, 0.5% perhaps. This prospect is not of value per se but only if the aim in the pursuit of screen-prompted diagnosis is to take advantage of its associated higher rate of curability relative to that associated with symptom-prompted diagnosis, perhaps a 40 percentage-point reduction in the overall rate of fatality (from 90% to 50%). Now, insofar as these two numbers are applicable, the instance of unit screening provides a 0.005 × 40% = 0.2% probability of curing (by early intervention) an already present though still asymptomatic lung cancer, fatal if left for symptom-prompted diagnosis. This probability has its ultimate relevance in its consequent gain in life expectancy (from early intervention, in lieu of late). The screenee’s current age etc. define life expectancy conditional on current absence, or cure, of lung cancer, perhaps 10 years longer than in the presence of lung cancer left for symptom-prompted diagnosis; and with this added input the gain in life expectancy would be 0.005 × 0.4 × 10 years = 0.02 year, resulting from a single repeat screen. In the particular context of screening for lung cancer, applied to apparently healthy people with the principal aim of preventing premature death, it is indeed reasonable to think of effectiveness in terms of added life-years per se, without having to deal with quality-adjusted life-years.

As for potential attainability of cost-effective screening, then, the very first point of contemplation might be that of how the screening, with whatever regimen, should be organized so as to minimize its unit (per unit screening) cost, that is, to maximize the efficiency of its implementation, to make it efficient. Beyond this, and in reference to this organizational framework, there is of course the challenge of designing the screening regimen itself with a view to maximization of its cost-effectiveness in various potential applications. And fi-
nally, the applications themselves, as to their indications within and across individuals, need to be designed with a view to attainment of cost-effectiveness.

In respect to a screening program thus designed for maximal cost-effectiveness, the need is to know or surmise the magnitude of its unit cost, C, and the magnitudes of the component parameters that jointly imply the effectiveness, E, of the unit screen, the latter expressed as the resulting gain in life expectancy. Then, C/E represents the perceived cost-effectiveness, usually expressed as the cost per life-year gained. Thus, if the unit cost of the screening (as organized) were taken to be, say, $200 and the effectiveness of its unit application the 0.02-year gain in life expectancy considered above, then the corresponding cost-effectiveness of the screening would be expressed by $200/0.02 = $10 000 as the cost per life-year gained. The smaller this cost is, the higher is the cost-effectiveness.

Screening, like any activity in health care, is said to be cost-effective if its cost-effectiveness is as high as is practically attainable for the purpose and, as such, high enough to justify the activity.

**How to think about screening-associated curability and its implications for mortality?**

Interest in the possibility of cost-effective screening for lung cancer derives, principally, from the near-complete incurability of this disease — its case fatality rate of some 90% — when left to be diagnosed on the prompting of symptoms (or an incidental finding in chest imaging).

The idea that the dismal, mere 10%, rate of curability in the absence of screening would be improved by screening is deduced from two premises, one diagnostic and the other interventional. The diagnostic premise is, naturally, that screening leads to diagnoses at less advanced pathologic stages of disease progression — including more common diagnoses in stage I. Just as naturally, the interventional premise is that curability is more common in the context of diagnoses at relatively early stages — especially in stage I.

This idea of enhanced curability in association with screen-prompted diagnoses, derived from those qualitative premises, is qualitative only; but any contemplation of cost-effectiveness requires a quantitative outlook. In the latter, the point of departure is the adoption of a scale, an operational one, in terms of which it is possible to express how “early” any given diagnosis of lung cancer is. The adopted scale might be the familiar ‘TNM scale’ of stages; it might be one of categories of size of the tumour; or it could be a combination of those two (beyond the size dichotomy in stage I). Given such a scale, one is concerned to know, or at least to surmise, for each of its categories the corresponding quantitative entries from the screen-diagnostic distribution and the curability function, respectively — the proportion, P, of screen-diagnosed (asymptomatic) cases falling in this category and the rate, R, of curability of diagnosed cases — asymptomatic — falling in this category. These quantities imply the overall rate of curability for screen-diagnosed cases (50% in the example above) — as the sum of the P, R, products over all of the categories. It may bear emphasis that while those diagnostic parameters (P) are peculiar to the screening regimen at issue, the interventional ones (R) presumably are not; and further, that it is reasonable to presume both sets of parameters to be independent of the risk profiles and life expectancies of the screenees.

The impact of screening on fatality rate — the complement of curability rate — depends not only on the curability of screen-detected cases, but also on that of cases detected when symptoms are exhibited before the next scheduled screening; these two rates are quite possibly different even when disease stage and tumour size are factored in. The overall rate of curability of cases diagnosed under screening, 45% perhaps (Table 1), is a weighted average of the overall rates of curability for the screen- and symptom-detected cases, with weights proportional to the respective frequencies of the two types of diagnosis under screening. The complement of this combined overall rate of curability — the fatality rate under screening — as a fraction of that prevailing in the absence of screening (55%/90% = 61% in the example above) implies (as its complement, 39% in the example) the percentage decline in lung-cancer fatality rate that the screening would provide for among the screenees.

This reduction in fatality rate translates to the same proportional decline in lung-cancer mortality rate among the screenees, though only after a lag time corresponding to

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Prompting of diagnosis</th>
<th>Proportion of cases diagnosed, %</th>
<th>Curability rate, %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>With screening</td>
<td>Without screening</td>
</tr>
<tr>
<td>Screening</td>
<td>Positive result of a screening test</td>
<td>80</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>Symptoms</td>
<td>20</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>Either</td>
<td>100</td>
<td>45</td>
</tr>
<tr>
<td>No screening</td>
<td>Symptoms</td>
<td>100</td>
<td>(45)</td>
</tr>
</tbody>
</table>
the maximum of the time from screen-prompted diagnosis (of very small malignancy) to death from the cancer (relatively slowly growing), if not cured; it thus is fully manifest in those who have been screened for, perhaps, 10 years or longer (and continue to be screened). For the screened and unscreened combined, the screening would provide for this decline multiplied by the proportion of cases of lung cancer that occur among the screenees. With only a small proportion screened, the latter decline is quite small; but this has no bearing on the cost-effectiveness of the screening, as both the cost and the benefit are confined to the screenees.

**How high might the cost-effectiveness be?**

At present, ideas about potentially cost-effective screening for lung cancer center on the use of **CT scanning**, on screening regimens pivoting on this type of screening test and involving CT in the further diagnostics as well; and more specifically, they now tend to center on the particular CT-based regimen devised by the clinical investigators in the pioneering ELCAP now underway.

My discussions with clinicians who have concerned themselves with this topic lead me to understand that with efficient organization, it should be possible to conduct ELCAP-type “screening” — i.e., pursuit of early diagnosis, rule-out or rule-in — for a unit cost well under US$200, the amount I used — quite conservatively — in the illustration above. In repeat screening, a positive test result is a matter of finding a new (newly detected) “noncalcified” nodule, one that has grown since the previous screen. Its ELCAP-type work-up begins with a course of antibiotics, followed by assessment of the result by re-imaging. This may show complete resolution and thus end the work-up. Otherwise, documented absence of further growth serves to obviate biopsies (CT-guided) — just as documentation of further growth, at an appropriate rate, serves to justify biopsy and to support possible cytologic diagnosis of malignancy.

Following that illustration further, could the curability rate for cases of lung cancer detected by ELCAP-type screening be as high as the 50% used in that illustration (Table 1)? At baseline in the ELCAP, 23 of the 31 diagnosed cases (all screen-diagnosed) were non-small-cell carcinomas of stage I; and it seems quite reasonable to presume that experience with repeat screening will show the proportion of non-small-cell cases of stage I to be about 70% among all screen-diagnosed cases. In the context of classical radiographic-and-cytologic screening the rate of curability of stage I cases — inherently truly malignant and not “overdiagnosed” — is about 70%. The CT counterpart of this — with almost all of the stage I cases actually

The advent of CT has recently provided major advances in respect to the screening test for lung cancer and its associated definitive diagnostics, while little has happened in the development of interventions to supplement resection of the tumour. Thus, the prospects for cost-effective screening for lung cancer suddenly are brighter than ever before, while in the absence of screening the disease remains as incurable as it has been in the past.

Overall, these considerations imply that CT-based screening for lung cancer, suitably specified, can be presumed to save lives at a cost lower than the US$10 000 per life-year saved in the illustration above. This level of cost-effectiveness is very well within the range of practice-acceptability, implying that suitably specified CT-based screening for lung cancer can be presumed to be quite cost-effective.

**Implications for policy**

The presentation above, concerning the economic and especially the medical underpinnings of rational policies on screening for lung cancer, focuses on the current context of available types of screening test, further diagnostics and interventions; and it bears emphasis that the advent of CT has just recently provided major advances in respect to the screening test and its associated other, definitive diagnostics, while little has happened in the development of interven-
tions to supplement resection of the tumour. Thus, the prospects for cost-effective screening for lung cancer suddenly are brighter than ever before, while in the absence of screening the disease remains as incurable as it has been in the past.

As for how bright the prospects now are, I have sketched my perception of this; and the very first implication of this perception, were it to be accepted and acted upon by policy-setting experts, would be that the still prevailing recommendations against whatever possible screening for lung cancer in the United States and Canada, different from that in Japan, would be rescinded. After this updating of the broadest policy — to allow for the possibility of cost-effective screening under suitable specifications — the authorities would set out to actually specify the respective regions of recommendation for, and policy neutrality. Each of these specifications would be guided by appreciation of the principal determinants of cost-effectiveness addressed above: the elements in the regimen’s unit cost and the determinants of the regimen’s effectiveness in application.

Policy-setting experts on this topic scarcely can defensively ignore the essence of my assessment of the current prospects as for cost-effective screening for lung cancer; and they especially cannot deny the fundamental fact that the advent of screening CT (helical, low radiation-dose) has created a new situation that calls for re-examination of policy. But if, even in this new situation, the policy-expert view is that cost-effective screening still cannot be envisioned with any combination of types of screening and screenee, brought together with whatever degree of attainable efficiency, this persistent official nihilism needs to be documented and made public, its “justification” included.

In the meantime, awkward encounters can occur in practice. People concerned with their prospects of getting to be diagnosed, in the absence of screening, with incurable lung cancer will continue to raise the question of screening for this disease in the hopes that early diagnosis and its consequent early intervention would avert such a possibility. Such a person might ask the doctor whether modern screening indeed provides for earlier diagnosis, and whether earlier diagnosis indeed now provides for more effective intervention. The doctor would have to give a confidently affirmative answer to each of these two questions, including in respect to pre-CT types of screening, not as a matter of personal belief but of general knowledge in medicine, and from this the concerned inquirer would deduce the idea that modern screening for lung cancer indeed can serve to avert death from this otherwise essentially incurable and near-uniformly fatal disease, the idea that modern screening for lung cancer (cum early intervention) indeed is known to be effective in preventing death from this dreaded disease. The doctor would have to reaffirm the correctness of the premises and acknowledge that the deduction from those is logically impeccable; and then, somewhat flustered, the doctor would have to proceed to point out that even the most modern and promising type of screening for lung cancer is not officially known to be effective, much less officially known to be cost-effective so as to justify its coverage by health insurance — even on the most compelling of indications and irrespective of how low the cost might be. Now the inquirer, in considerable disbelief, would be very curious to learn what it is in official medicine that can override logical deduction from securely agreed-upon premises, leading to an official stance of agnosticism or nihilism even. The doctor, already overtly flustered, would need to point out that, in official medicine at present, formally correct conclusions deduced from materially correct premises can be questioned, and indeed totally ignored, so long as there is lack of direct evidence, notably from RCTs — including here, despite their irrelevance for the diagnostic subissue and their ethical inadmissibility as well as scientific superfluousness for the interventional subissue.

Perhaps the most fundamental policy issue in this context is the degree to which there is to be genuine and uncompromising commitment to reason in lieu of the currently pervasive adherence to the doctrine of RCTism. With liberation from this doctrine, limited practice (as to indications especially) could already be recommended, and research resources could be directed to expansions of and innovations in the recommendations.

Policy updates, such as they are

At the time of this writing, both the American Cancer Society and the US National Cancer Institute are distancing themselves, ever so slightly, from their existing recommendations against screening for lung cancer.

The American Cancer Society now expresses itself as follows:

In spite of the limitations of the existing data, it is generally accepted that lung cancer screening is not effective, whereas it would be more appropriate to regard the current evidence-based situation as one in which there are insufficient data to recommend for or against lung cancer screening. ... [The recommendations from the Verase [sic] meeting have prompted the ACS to initiate a process for reconsidering the current advice, potentially stressing both the limits of the existing data, as well as the paradoxical findings of trial results versus case-finding series.
So, the American Cancer Society now expresses the stance of wholesale agnosticism only, still not knowing any available type of screening for lung cancer on any type of screenee to have been, or even now to be, at all effective in preventing death from this disease (through early intervention); and further, were the American Cancer Society to come upon positive “trial” (RCT) results on effectiveness, it would take this qualitative evidence to be a sufficient basis to recommend for screening, without regard to any specifics — in sharp contrast to the analytic, quantitative, cost-effectiveness outlook described above, which is specific to particular types of screening and screenee. Direct evidence, even if only qualitative and only subjectively (mis)interpreted, rules the “evidence-based situation,” stifling reason — as in Evidence-Based Medicine practiced in place of reason-and-knowledge-based medicine.12

The National Cancer Institute, by contrast, does not appear to be about to reconsider its nihilistic advice. But it, too, appears now to be agnostic about the effectiveness of modern screening for lung cancer, as it is just now seriously considering the investment of hundreds of millions of US dollars in an RCT that, perhaps no earlier than 15 years from now, would serve to test a qualitative — and knowingly false — “hypothesis” of no difference in effectiveness between two turn-of-the-century regimens of lung cancer screening: upon “informed” consent, misinformed people would be randomized to knowably superior (CT) and inferior (x-ray) screenings — both presumably obsolete in 2015, to say nothing about the policy-irrelevance of qualitative ideas even in respect to state-of-the-art approaches. RCTism at the extreme of its irrationality!

Canadian authorities are yet to express themselves in the new situation that the ELCAP has, however inadvertently, brought about. How timely, rational and knowledge-based their updated recommendations will be, this remains to be seen. But they will be anything but consequential, as countless lives hang in the balance, within Canada and, especially, in the world at large. Ex Canada lux?

Dr. Miettinen is with the Department of Epidemiology and Biostatistics and the Department of Medicine, Faculty of Medicine, McGill University, Montreal, Que., and also with the Department of Medicine, Cornell University, New York City.

Competing interests: Dr. Miettinen is an investigator in the ELCAP.

References

Correspondence to: Dr. Olli S. Miettinen, Department of Epidemiology and Biostatistics, Faculty of Medicine, McGill University, 1020 Pine Ave. West, Montreal QC H3A 1A2; fax 514 398-4503; osm@epid.lan.mcgill.ca

CMAJ Essay Prizes
Deadline: Dec. 15, 2000

CMAJ will award prizes for the best essays on any health-related subject submitted during calendar year 2000. A $2000 prize will be awarded for the best entry submitted by a medical student or resident. There is also a $2000 prize for the best entry submitted by any author. These new contests replace the Logie Medical Ethics Essay Contest for medical students.

We are looking for reflective essays of up to 1500 words. Manuscripts must be original and must be submitted only to CMAJ. Winners will be selected by a committee appointed from the CMAJ Editorial Board. Winning entries will be selected based upon originality, quality of writing and relevance to health or health care. To win, a manuscript must be suitable for publication. If suitable entries are not received, prizes will not be awarded. All papers submitted will be considered for publication in CMAJ.

Authors should submit their papers with a covering letter stating that they wish the manuscript to be considered for the essay prize, and should indicate their status regarding training. Send entries and queries to: Dr. John Hoey, 1867 Alta Vista Dr., Ottawa ON K1G 3Y6; hoeyj@cma.ca