

## 特別報告 肺がん検診の有効性評価：現行検診と胸部 CT 検診

佐川元保<sup>1</sup>・杉田 真<sup>1</sup>・佐久間勉<sup>1</sup>

**要旨** 肺がん検診の有効性評価に関しては、1970～1980年代初頭の検診・診断・治療水準を評価した無作為化比較試験では有効性を証明できなかったが、1990年代の水準を評価した4つの症例対照研究では検診受診により30～60%救命可能なことが判明した。しかし、全国的に見れば、精検受診率が低い自治体もある。県の成人病検診管理指導協議会が中心となって、精度管理の水準を地区ごとに評価し公表することが必要である。胸部CT検診では、肺癌発見率と生存率は驚異的に高いが、シミュレーションによれば発見数は急速に低下するはずであり、多数の発見が続くならば、それは「死に至らない肺癌」すなわち Overdiagnosis bias による可能性が示唆された。剖検例のCT所見の検討によると、死亡前に診断されず剖検で診断された肺癌が相当な比率に上り、剖検でチェックされずCTでは指摘できた肺腫瘍はさらに多数であり、Overdiagnosis bias は胸部CT検診の結果に重大な影響を与えている可能性がある。真に有効性を証明するためには無作為化比較試験を行うのが望ましい。(肺癌・2002;42:846-850)

**索引用語** 肺がん検診, 有効性, 精度管理, 胸部CT検診, オーバーダイアグノーシス・バイアス

## The Efficacy of Lung Cancer Screening: The Present Japanese Screening System and Thoracic CT Screening

Motoyasu Sagawa<sup>1</sup>; Makoto Sugita<sup>1</sup>; Tsutomu Sakuma<sup>1</sup>

**ABSTRACT** Although the results of randomized controlled trials evaluating the efficacy of lung cancer screening in the 1970s early 1980s failed to show that lung cancer mortality would decrease by screening, four case-control studies evaluating the efficacy of lung cancer screening in the 1990s were recently reported from Japan. These indicated that screening would decrease the risk for lung cancer deaths by 30-60%. However, quality control in the screening system is not enough in some parts of Japan. Prefectural governments should evaluate the quality of screening systems in each region and report the results. Preliminary reports concerning lung cancer screening using thoracic CT revealed that not only the detection rate of lung cancer but also the survival rate of detected lung cancer patients were very high. However, overdiagnosis bias must affect the results, and a randomized controlled trial should be conducted to evaluate the true efficacy of lung cancer screening using thoracic CT. (JJLC. 2002;42:846-850)

**KEY WORDS** Lung cancer screening, Efficacy, Quality control, Thoracic CT screening, Overdiagnosis bias

### はじめに

老人保健法が施行されて以来、肺がん検診は本邦で広く行われてきたが、その有効性、すなわち「肺がん検診を受診すると本当に肺がんで死亡する危険性が減るのか」という点に関しては、ごく最近まで強く疑問が投げ掛けられていた。本稿では、現行検診の有効性の評価に

関するこれまでの流れと、胸部CT検診における有効性評価の現状に関して述べる。

### 検診の有効性評価に関わるバイアス

肺がん検診の有効性に関して、もっとも単純な検診は、検診発見例と非検診発見例の生存率の比較である。以前からこのような研究は数多くなされてきており、その多

<sup>1</sup> 金沢医科大学呼吸器外科。

<sup>1</sup>Department of Thoracic Surgery, Kanazawa Medical Univer-

sity, Japan.

© 2002 The Japan Lung Cancer Society

くは検診発見群の予後が有意差をもって良好であることを示すものであった。しかしながら、現在においては、そのような結果だけでは有効性を証明したことにはならないと考えられている。それは、検診発見群と非発見群の予後を比較する際に、以下のようないくつかのバイアスが存在しているからである。

検診で発見された癌は、症状が出現する前に発見されたのであるから、症状が出現してから診断される「症状発見群」に比較して、症状が発現するまでの期間分長生きする、というバイアスが Lead time bias である。症状が発現するまでの期間を Lead time と呼び、その期間だけのバイアスがある、ということになる。検診で発見されて6年で死亡した人が、実は検診を受けてから症状が出るまで2年かかるような癌であった場合には、検診を受けなかった場合は発見されてから4年で死亡することになる。したがって、検診による救命効果も延命効果も存在しなくとも「症状群は診断後4年で死亡」検診群は診断後6年で死亡」となり、生存率には見かけ上大きな差異が生じることになる。

検診で発見される癌は、検診の時期まで症状が出現しないような癌であり、もし、もっと速い速度で増大するような場合は、検診の合間に症状で発見されることになる。したがって、検診ではゆっくり増大する癌が多く発見され、そのため発見後の進行の速度も比較的ゆっくりなため、見かけ上、検診発見群は症状発見群よりも長生きする、というバイアスが Length bias である。たとえば、毎年検診を受けていても症状で発見され、治療の甲斐なく死亡する人は存在するが、このような人は毎年検診を受けていたのであるから、検診の益は全くなかったわけである。しかしながら、生存率の比較では、このような人を「症状発見群」に入れてしまうことになり、結果として検診に有利なバイアスがかかることになる。

ゆっくり増大するため、もし検診がなければ、症状が出現しないうちに最終的に他の疾患で死亡するような癌を、検診があるために見つけてしまうことが Overdiagnosis bias である。実際、このようなものは、検診をしなければ見つからずに終わっていたはずのものであり、見つけなくとも実害のないものである。このバイアスは、小さい癌、早期の癌を見つけようとすればするほど問題になってくる。通常は見つからないで終わるものを見つけていることになるので、発見率は極めて高くなる。また、その癌では死亡しないので、当然生存率も高くなる。したがって、生存率や発見率がいくら高くても、それらが Overdiagnosis bias のためであれば、真の肺癌死亡の減少に結びつかない。

検診を受診するような人は、生活習慣上も健康に留意しており、その結果当該疾患に罹患あるいは死亡する確率が低い、すなわち集検受診者の死亡率が低くてもそれ

は集検の効果とは言えない、というバイアスが Self-selection bias である。他の3つのバイアスが、どちらかと言えば生存率に関係するのに比べて、このバイアスは死亡率にも関係する点で重要である。

## これまでの有効性評価の成績

前述した各種のバイアスのために、単純な生存率の比較だけでは、効果の有無を云々することは不可能である。現在、検診の有効性を評価する際に信頼されている標準的な方法としては、無作為化比較試験、コホート研究、症例対照研究などが挙げられる。これらの方法で行われた現在までの成績は久道班報告書<sup>1,2</sup> に詳述されているが、ここではそれを概説する。

### 1. 初期の研究 (1960~1970年代の検診を評価)

肺癌の早期発見・早期治療を目的として無症状の成人を対象に胸部X線写真を撮影することの効果の評価するために、いくつかの介入研究が行なわれたが、その結果は死亡減少効果に関して否定的であった。

### 2. 欧米での無作為化比較試験および症例対照研究 (1970~1980年代初頭の検診を評価)

米国NCIは喀痰細胞診が肺癌の早期発見に有用であることが判明したことを踏まえて、無症状の成人に対して行なう喀痰細胞診+胸部X線写真が肺癌死亡減少に寄与し得るかどうかに関する無作為化比較試験をMayo Clinicを始めとする3施設で行った。その結果は、「肺がん検診が肺癌死亡を減少させるとは証明できない」というもので、チェコでの無作為化比較試験・東ドイツでの症例対照研究における同様の結論と相まって、その後の世界の潮流は、肺がん検診を社会的な施策として取り入れない方向へ進んだ。これらの研究に関しては、多くの研究者からさまざまな解釈上の問題点が指摘されているが<sup>3,4</sup> それに関しては別著<sup>5</sup>を参照されたい。特に重要な点として、これらの結果は20~30年前の診断治療水準(CTですら標準的ではない)でのものであることは銘記すべきであろう。

### 3. 本邦における初期の症例対照研究 (1980~1990年代初頭の検診を評価)

本邦では、結核予防法により撮影された胸部X線写真を利用する形で肺がん検診が始まった。喀痰細胞診の導入もなされ、早期肺癌が続々と発見されるようになり、厚生省成毛班による初めての症例対照研究が行われた。50市町村から273セットの症例及び対照の集積があり、オッズ比(集検を受診した人が受診しない人に比較してどの程度肺癌死亡しやすいか)は0.72と算出された。すなわち、28%が集検で救命されていると考えられた。有意差にはわずかに足りなかったが、死亡減少効果を示唆する結果であった<sup>6</sup>。続いて、金子班において個別検診に関する症例対照研究が行なわれ、オッズ比0.531という

値で、有意な死亡減少効果を示した。<sup>7</sup>

#### 4. 本邦における最近の症例対照研究（1990年代の検診を評価）

厚生省藤村班では、1997年から全国4地区（岡山、新潟、群馬、宮城）において同時に症例対照研究を遂行した。詳細は別著<sup>8-12</sup>を参照されたいが、それぞれの地区では、ケースおよびコントロールを数百例から数千例調査した結果、診断の12ヶ月前以内に検診受診することの肺癌死亡に関するオッズ比は0.4~0.7程度で、3地区では有意な死亡リスクの減少効果を示した。このことは、現行の方法でも肺がん検診受診により30~60%救命可能なことを示している。一方、同時に行った検診受診からの期間別の検討では、診断前の12~24ヶ月に検診を受診しても肺癌死亡に関するオッズ比は低下しないことも判明した。検診の効果が1年しか持続しないという点で、肺癌は治癒させにくい癌だということができよう。また、1回の見逃しが治癒の機会を失わせる可能性についても示唆している。

#### 現行検診に関する問題点

藤村班の結果は肺がん検診に死亡減少効果があることの証拠となるものであったが、実際にすべての都道府県で行われている肺がん検診が同等の効果を有するものであろうか？

「がん検診の精度評価に関する手引き」<sup>13</sup>の中で久道、辻らがまとめた肺がん検診の各都道府県別精検受診率によれば、精検受診率が60%前後の自治体もあり、このような状況ではどのように精密な読影や詳細な精密検査を行おうが、疑わしいものの4割は無条件に見逃していることになるわけで、肺癌死亡の減少を期待することは難しいと思われる。藤村班の研究を行った4地区では、いずれも精検受診率は80~90%あり、検診により肺癌死亡を減少させるためには、読影や検査を綿密に行うのみでなく、検診のシステム自体を緻密にかつ地道に穴埋めしていく作業が必要である。肺がん検診はスクリーニングを行うことで完結するのではなく、その後の精密検査・治療・追跡・受診勧奨等が有機的なシステムとして動かなければ肺癌死亡率の低下に繋がることは難しい。

一方、平成10年度から肺癌検診にかかわる国庫補助が一般財源化され、各自治体の実状に即した検診が可能となった。この変更により、安価ではあるが質の悪い集検がはびこることも懸念されている。スクリーニングのみを安価で提供し、その後には関知しない、という立場の企業が台頭してくるであろう。これを機に、ダンピングに走るのではなく、精度管理にかかわる事項を広く情報公開することによって、「悪貨を駆逐する」ことが肝要であろう。

具体的には、都道府県の成人病検診管理指導協議会が

それに関して指導的役割を果たすべきであろう。この協議会は、これまではむしろ組織としての活動はほとんどなされていないのが多くの県の実態であったが、検診団体が乱立して、がん検診の精度管理の水準の維持が危ぶまれている今こそ、強力なリーダーシップのもとで精度管理の水準を検診団体ごとに評価し公表することが必要とされており、構成メンバーの布陣から言ってもそれが可能な組織である。法的・予算的・人員の基盤が不十分であることには国の責任が大きい、充分とは言えないにしても若干の強制力や事務費・通信費等は確保されているのであるから、その中でも最低限の調査は可能である。調査すべき内容に関して、「がん検診の精度評価に関する手引き」<sup>13</sup>の中に詳細なチェックリストが公表されているので、これをベースとしながら「老人保健法による肺がん検診マニュアル」<sup>14</sup>または「肺癌取扱い規約」<sup>15</sup>の中の「集計表」を埋めるために必要な情報が整理されているかどうかを加えることでほぼ充分であろう。さらに具体的には、別稿にて調査様式のひな形を提示し調査対象に関して詳述したのでそちらを参照されたい。それらを参考にすれば、多忙な成人病検診管理指導協議会のメンバーでも、県の事務系に調査方法を指示し1年に1度調査させることは容易である。住民が自分たちの地区における検診を客観的に評価できるような情報を公開することこそが「情報公開」であり、それに基づいて、コストとのバランスも含めた自治体単位の政策の決定がなされるべきであろう。

#### CT 検診の有効性評価

胸部CT検診では、その高い肺癌発見率と発見肺癌の生存率は、まさに驚異的であり、今後の肺がん検診のツールとして大きな可能性を有している<sup>16-18</sup>。しかしながら、現在までのデータでは、未だ有効性が確認されているとは言えない。それには、バイアスなどがより大きく関わってきている可能性があるからである。

まず、発見率に関してであるが、CT検診は現行法の何年か分を1年で発見していると考えられるとは言え、報告されるほど発見率が高くなるものなのであろうか。前述したOverdiagnosis Biasによる見掛け上の発見率の高さではないのか。その点を明らかにするために、CT検診が現行法の数年分を1年で発見可能と仮定して、経年受診させたときのモデルを用いたシミュレーションを行なった。詳細は別稿<sup>19</sup>を参照されたいが、初年度に現行法の5年分を発見可能と仮定した場合、2年目に発見可能なのは2年目の新規発生例と前年の見逃し例を合わせた数ということになり、見逃し率20%とすれば、3年目にはプラトーの1.1倍程度まで肺癌発見数は低下すると考えられる。見逃し率50%としても、5年目には1.1倍以下まで下がるはずである。さらに初年度に現行法の10

年分を発見可能と仮定した場合でも、見逃し率 20% とすれば 4 年目にはプラトーの 1.1 倍以下まで肺癌発見数は低下し、見逃し率 50% としても 6 年目には 1.1 倍程度まで下がる。このように、見逃し率を高めに設定してさえ、プラトーの 1.1 倍程度にまで発見数が低下するまでには、数年しかかからないはずである。そして、別に行った現行検診の感度分析の結果からは、このプラトーの値は現行検診での肺癌発見率の 1.3 倍程度にしか増えることはなく、それを超えて多数例の発見が続くようであれば、それは「死に至らない肺癌」すなわち Overdiagnosis bias による発見例の増加である可能性がある。

また、Dammas ら<sup>20</sup> による剖検例の CT 所見の検討によると、剖検例 1047 例中、死亡前 2 ヶ月以内に胸部 CT を受検した 187 例の CT 所見を検討し、15 mm 以下で 10 個以下の陰影のあるものが 28 例存在した。そのうち 9 例は剖検でチェックされておらず、19 例が剖検時にチェックされていた。19 例中 2 例が死亡前には診断されなかった肺原発の扁平上皮癌であった、ということである。187 例中 2 例が未診断の肺癌ということであると、10 万対比 1000 を越えてしまう。したがって、その程度の死亡に関係しない肺癌 (Overdiagnosis bias) が隠れていて、胸部 CT 検診ではそれらを余分に発見しているのかもしれない。1047 例中と考えても 10 万対比 200 位になり、また、9 例は剖検でチェックもされなかったということも考えると、Overdiagnosis bias は胸部 CT 検診の結果に重大な影響を与えている可能性がある、と言わざるを得ない。

### 胸部 CT 検診の有効性評価に関する研究

前述したように胸部 CT 検診に真の意味での有効性があるのかどうかを知ることはきわめて重要である。現在の枠組みの中では、厚生労働省の鈴木班で大規模なコホート研究が動き始めたところであり<sup>21</sup>、現在 CT 検診を行なっている施設は、その研究に参加されたい。結果は 5 年後か 10 年後になるだろうが、positive data が出れば、有効性 (効果) に関する大きな evidence の一つとなるであろう。

しかしながら、コホート研究は無作為化比較試験 (RCT) よりも証拠の質が低いとされているのも事実である。もし日本のコホート研究がうまくいっても、近い将来に他国で行われた RCT の結果が「効果なし」となれば、どんなに向こうの検診・精査・治療の水準が低くても、向こうの方が勝つのである。日本で RCT ができない最も大きな要因は研究費用が無いことであるが、無いが無いなりに、研究者のまじめさと能力、それにこれまで培った地域とのパイプを用いて、RCT を行なう計画を立案した。詳細は別稿<sup>22</sup> に述べるので、それを参照されたいが、概略を述べると、住民検診受診者から希望者をリクルー

トし、十分な Informed Consent を得たうえで、数ヶ月後に希望者の半数に CT 検診車のレンタルを用いた胸部 CT 検診を 1 回のみ受診して頂き、費用は研究費を用いるというものである。1 地区あたり 1000 人ずつ、総計 20 地区で 20000 人を予定し 10 年間フォローする。この計画は 1 地区当たりの作業負担は少なく、研究者個人の努力でもカバー可能である。研究費も 1 地区あたり 500 万円程度、総額 1 億円程度で可能であり、また、地区を徐々に増やして行って、最終的に 20 地区に達するまで行う、といった計画も可能である。有効である証拠がないまま胸部 CT 検診が中途半端に広まって収拾がつかなくなる前に、世界中からの評価に耐える研究を行う必要がある。

### REFERENCES

1. 久道 茂, 編. がん検診の有効性評価に関する研究班報告書. 東京: 日本公衆衛生協会; 1998.
2. 久道 茂, 編. 新たながん検診手法の有効性の評価報告書. 東京: 日本公衆衛生協会; 2001.
3. Fontana RS, Sanderson DR, Woolner LB, et al. Screening for lung cancer: a critique of the Mayo Lung Project. *Cancer*. 1991;67:1155-1164.
4. Strauss GM, Gleason RE, Sugarbaker DJ. Screening for lung cancer. *Chest*. 1997;111:754-768.
5. 佐川元保, 中山富雄, 塚田裕子, 他. 肺癌検診の有効性評価に関する新しい知見. 日胸. 1999;58:S20-S24.
6. Sobue T, Suzuki T, Naruke T, et al. A case control study for evaluating lung cancer screening in Japan. *Int J Cancer*. 1992;50:230-237.
7. Okamoto T, Suzuki T, Hasegawa H, et al. Evaluation of a clinic-based screening program for lung cancer with a case-control design in Kanagawa, Japan. *Lung Cancer*. 1999;25:77-85.
8. 佐川元保, 中山富雄, 塚田裕子, 他. 肺がん検診の有効性評価 厚生省藤村班での 4 つの症例対照研究. 肺癌. 2001;41:637-642.
9. Nishii K, Ueoka H, Kiura K, et al. A case-control study of lung cancer screening in Okayama Prefecture, Japan. *Lung Cancer*. 2001;34:325-332.
10. Tsukada H, Kurita Y, Yokoyama A, et al. An evaluation of screening for lung cancer in Niigata Prefecture, Japan: a population-based case-control study. *Brit J Cancer*. 2001; 85:1326-1387.
11. Sagawa M, Tsubono Y, Saito Y, et al. A case-control study for evaluating the efficacy of mass screening program for lung cancer in Miyagi Prefecture, Japan. *Cancer*. 2001;92: 588-594.
12. Nakayama T, Baba T, Suzuki T, et al. An evaluation of chest X-ray screening for lung cancer in Gunma Prefecture, Japan: a population-based case-control study. *Eur J Cancer*. 2002;38:1380-1387.
13. 久道 茂, 編. 平成 9 年度厚生省「成人病検診管理指導協議会のあり方に関する調査研究」研究報告書: がん検診の精度評価に関する手引き. 東京: 日本公衆衛生協会; 1998.
14. 厚生省老人保健福祉部老人保健課. 老人保健法による肺がん検診マニュアル. 東京: 日本医事新報社; 1992.

- 15 . 日本肺癌学会集団検診委員会 . 肺癌集検の手引き . 日本肺癌学会 , 編 . 肺癌取扱い規約 . 第 5 版 . 東京 : 金原出版 ; 1999:159-177
- 16 . 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. *Brit J Cancer*. 2001;84:25-32.
- 17 . Kaneko M, Kusumoto M, Kobayashi T, et al. Computed tomography screening for lung carcinoma in Japan. *Cancer*. 2000;89:2485-2488.
- 18 . Henschke CI. Early lung cancer action project: overall design and findings from baseline screening. *Cancer*. 2000;89:2474-2482.
- 19 . 佐川元保, 斎藤泰紀, 佐藤雅美, 他 . 肺癌の疫学 最近の知見 . 総合臨床 . 2001;50:2217-2222.
- 20 . Dammas S, Patz EF, Goodman PC. Identification of small lung nodules at autopsy: implications for lung cancer screening and overdiagnosis bias. *Lung Cancer*. 2001;33: 11-16.
- 21 . らせん CT による肺がん検診の肺がん死亡減少効果検討のためのコホート研究計画書 . 鈴木隆一郎, 編 . 肺がん検診における高速らせん CT 法の効果評価研究報告書 . 東京 : 日本公衆衛生協会 ; 2000.
- 22 . 佐川元保, 杉田 真, 佐久間勉 . 胸部 CT による肺がん検診の有効性評価に関する無作為化比較試験 . 胸部 CT 検診 . 2002;in press.