

有効性評価に基づく肺がん検診ガイドライン更新版 公開フォーラム

議事概要

日時：2025 年 2 月 19 日(水) 14:00～16:40

場所：オンライン会議

パネリスト（順不同・敬称略）

中山 富雄	国立がん研究センター
濱島 ちさと	帝京大学
細野 覚代	国立がん研究センター
芦澤 和人	長崎大学大学院
桜田 晃	みやぎ県南中核病院

発言者（順不同・敬称略）

佐川 元保	東北医科薬科大学
八重樫 牧人	千葉西総合病院
斎藤 博	青森県立中央病院

1. 更新版の概要について

中山

本日の有効性評価に基づく肺がん検診ガイドライン更新版の公開フォーラムにあたり、更新版の作成方法、概要について簡単にお話しておきたい。

ガイドライン作成にあたっては、まず文献レビュー委員会がエビデンスレポートを作成し、それを基に幅広い分野の方々7名から成るガイドライン作成委員会が臨床にどう生かせるかを検討している。判断の基準としては、死亡率減少効果という利益と証拠のレベルの信頼度と不利益の程度を合わせて勘案し、推奨グレード A、C、D、I に分けている。

A: 利益あり、信頼度・不利益が中、対策型検診及び任意型検診両方で推奨

C: 利益あり、信頼度・不利益が小、対策型では推奨せず、任意型では条件付きで実施

D: 利益なし、信頼度なし、不利益が大、対策型検診及び任意型検診両方で推奨せず

I: 利益・信頼度は不明、不利益あり、C と同様。

肺がん検診においては対象を重喫煙者と非喫煙者と低喫煙者に分けていて、検診方法は対象に応じて低線量 CT、胸部 X 線、喀痰細胞診を実施している。重喫煙者にのみ認められる肺門部肺がんは喀痰細胞診での早期発見が可能、非・低喫煙者にも認められる末梢性肺がんには喀痰細胞診は有効ではない。低線量 CT 検診は電流を減らして被爆の軽減を図った撮影手法で、一定のトレーニングは必要だがすべての CT 撮影装置で撮影可能。

以上を基本的な概要としてご理解いただきたい。

2. 肺がん検診エビデンスレポート 2024 年度について

濱島

今回のエビデンスレポートにおいては 1990 年～2010 年代に行われた大型 RCT を中心に、幾つかのコホート研究等を新たなガイドライン作成のための評価材料としている。

重喫煙者に対する低線量 CT については NLST、NELSON の大型 RCT をはじめとした欧米の研究によって重喫煙者での研究が進んでおり、有意な死亡率減少効果が出ている。ただ過剰診断とインシデントアル・ファインディングの事例が多いことが問題。軽喫煙者・非喫煙者に対する低線量 CT の死亡率減少効果については日立コホート研究があるが、証拠は限定的である。

胸部 X 線については PLCO の大型 RCT を主な判断材料とした。論文の最終結論は 13 年の追跡終了時点で死亡率減少効果はなしとしているが、6 年時点で有意な死亡率減少効果が確認できることをもって文献レビュー委員会では 6 年時の結果を最終的な結論とした。参考研究として 1960 年代の

カイザーのRCTでは11年間の追跡で死亡率減少効果はないと結論付けているが、国内の症例対照研究においては死亡率減少効果を認めている。ただ、PLCOの追跡期間は最初から厳密に決まっていたのではなく、NLSTの報告とあわせて公表するために追跡期間13年で最終結果として報告したという経緯がある。重喫煙者について喀痰細胞診の併用についてはジョンズホプキンスとメモリアルスローンケタリングの研究において死亡率減少効果の傾向は認められるものの有意な結果は示されていないが、サブ解析をすると50パッキイヤー以上の重喫煙者には効果が確認されている。よく取り上げられるチェコとメーヨークリニックの研究は頻度の比較なので、有効性評価の検討材料としては不適。国内の症例対照研究は日本の肺がん検診システムを評価しておりメタアナリシスで有意な死亡率減少効果が見られるが、X線検査に喀痰を追加した上乗せ効果がどのぐらいあるかの評価は得られていない。

証拠のレベルとして文献レビュー委員会の結論は、低線量CTで重喫煙者については、科学的根拠あり、証拠の信頼性・不利益は中等度、利益は大。非・軽喫煙者については科学的根拠未定。胸部X線については、科学的根拠あり、証拠の信頼性は中等度、利益はそれほど大きくなく、不利益は中等度。喀痰細胞診併用については、科学的根拠なし、証拠の信頼性は中等度、利益は小、不利益は大とした。

3. 有効性評価に基づく肺がん検診ガイドライン更新版について

細野

今回、肺がん検診更新版を作成することになった背景には近年RCTの結果が複数報告されたことや日本の喫煙状況の変化等がある。ガイドライン作成委員会では文献レビュー委員会から提示された科学的根拠に基づいて議論し、推奨グレードを決定した。

低線量CTについては、重喫煙者については過剰診断や被爆の不利益があるものの推奨グレードはA、対象年齢50～74歳、検診間隔1年。非・軽喫煙者については、利益は不明、不利益は重喫煙者と同等以上なので推奨グレードはI。

胸部X線についてはPLCOの研究での減弱効果を考慮した6年時点の結果を採用し、国内の症例対象研究と合わせて推奨グレードA。対象年齢40～79歳、検診間隔1年。

喀痰細胞診については死亡率減少効果の上乗せを示す十分な科学的根拠がなく、喀痰細胞診のみによるがん発見数が大幅に減少しているということで推奨グレードDとした。

4. 質疑応答・パブリックコメントへの回答

指定発言

芦澤

日本肺癌学会肺がん検診委員会から 2022 年の肺がん検診ガイドラインを出していて、推奨の総括表を出しているのでご紹介したい。

低線量 CT については、重喫煙者は推奨レベル A、非・軽喫煙者は推奨レベル I ということで今回のガイドライン更新版と一致している。

胸部 X 線については危険群には喀痰細胞診を併用することも含めて推奨グレード A。こちらの更新版では胸部 X 線と低線量 CT の違いについて、重喫煙者に対しては低線量 CT の利益が大きいとの記載があって運用面にも配慮されているが、低線量 CT は被曝や読影医師の確保等、さまざまな課題があるので実装においては課題の解決が必要と感じている。

桜田

喀痰細胞診について意見を述べさせていただきたい。

2006 年のガイドラインでは高危険群に対する喀痰細胞診併用療法の文献を有効な根拠と判断して推奨グレードを B としている。今回の更新版は 2 つの RCT の統合解析だが、層別解析で 50 パックイヤー群がぎりぎりネガティブというだけで前回と正反対の判定としているのはかなり違和感があるので、前回同様、国内の症例対照研究も評価の対象とすべき。

発見がんの減少については精度管理の問題、コロナ禍による受診控え等、要因はいろいろ考えられるので慎重に検討する必要があるし、宮城県においては減っているが、減り続けているわけではなく安定して一定の発見率を有している。

論文の結論ではない統計データはエビデンスとは別に議論すべきだと思うし、それを判断材料とするなら事前に手順を決めておいて解釈のバイアスを最大限除くべき。また喀痰細胞診を終了とするのなら、これは他の検診にも大きく影響することなので、予め終了条件を決めておいた上で幅広い議論をして進める必要がある。

胸部 X 線と喀痰細胞診の併用に対する推奨は、エビデンスに基づいて A または C として、有効性を担保する条件を付すのが適当ではないかと考える。

【ディスカッション】

中山

チャットで法定健診との位置づけについて質問が来ているが、きょうはあくまで有効性評価に基づくガイドラインの内容のみで、この先の運用についてはずっと先にならないとわからない話なので、きょうお話しすることはできない。

まずは低線量 CT についてご意見をいただきたい。

佐川

濱島先生のスライドで Patz の過剰診断率、BAC で 17%というのは誤りだと思うのでそこは訂正したほうがいい。

濱島

確認して修正する。

中山

他に何かご意見があればお願いしたい。ないようであれば、低線量 CT については異論なしということで、次に胸部 X 線と喀痰細胞診についてご意見をいただきたい。

佐川

まず胸部 X 線について、PLCO で論文の著者がネガティブだと結論付けたものをポジティブの証拠だとするのはやはり無理。かろうじて言えることは、わずかながらも死亡率減少効果が認められるかもしれないという程度。著者がネガティブだと言っているものを、ポジティブだから A として推奨すると言われたら著者は怒りだすようなレベルの話。

濱島

PLCO の解釈については佐川先生のレターも含め当初から疑念のコメントが出ている。他のレポートでも 6 年目を解析し直したところはあるが、最終的に効果なしという判断になっている。PLCO は国内で行われた症例対照研究に比べて明らかに（エビデンスレベルが）優位な研究なので、症例対照研究で有意であってもそれのみで判断はできない。ただサジョンタイムを考慮すると PLCO の 6 年目での判断が妥当だろうと解釈を少し変えて判断している。論文の通りに解釈しなければならないということではないと思う。

佐川

PLCO の結果に国内の症例対照研究を加えたと言うが。

濱島

加えたとは言わない。PLCO が出たからにはそちらの判断が優先されるのは当然なので、それがネガティブなら症例対照研究の結果は判断材料にはならない。ただ 13 年目の判断が非常にポリティカルな事情だったので、そちらは採用せず 6 年目の結果で判断するのは適正と考えている。以前に RCT はなかったわけだから。

佐川

メイヨーの RCT があった。

濱島

メイヨーは頻度の研究。今回に関連するもの（検診の有無で有効性評価を判断した RCT）としては過去にはカイザーの研究しかないが、小研究であったため（検索できず）以前にガイドラインでは見ていなかった。今回は PLCO の 6 年目の結果と国内の症例対照研究の結果の両方を勘案して、X 線に関しては効果があると結論を出している。

佐川

メイヨーが頻度の差になってしまったのはあくまで結果論。

濱島

違う。最初から頻度の研究としてデザインされている。

佐川

そうじゃない。年に 3 回強力に勧奨する群に対して、片群には一応声をかけて放っておいたために結果的に頻度の差になっただけで、デザインとしては違う。

濱島

海外のエビデンスレポートでは高頻度介入群と低頻度非介入群を分けて評価しているデザインに分類されている。

佐川

PLCO に関しては海外を無視してポジティブと言っているのだから恣意的ではないか。海外のガイドラインでポジティブと言っているところはあるのか。

濱島

それはない。

佐川

ないでしょう。あるわけない。言うとしたら「ネガティブと判断しているけど、その判断には疑義もある」ぐらい。無理やりポジティブだと言って、日本の症例対照研究はあんまりあてにならない、みたいな言い方になっている。

濱島

あてにならないのではなくて、前回までは RCT がカイザーしかなかったから。

佐川

PLCO の RCT が出て X 線はネガティブになったと言うなら論理的には正しい。

濱島

ネガティブになった場合、日本の症例対照研究というのは CXR の根拠にならない。

佐川

ネガティブでいいのではないか。X線をネガティブとするなら論理的に整合性がある。

濱島

確かにガイドラインでは PLCO をポジティブと取ってないものがほとんどだが、先生の出されたレターと類似の見解も結構ある。

佐川

それはあくまでも疑義があるというだけであって、研究の外にいる人間が「あれはポジティブだ」とはとても言えないだろう。

濱島

そうだろうか。評価時点の 13 年というのは非常に恣意的に決められている。

佐川

先生の言っていることはわかるが、もし世界がポジティブだと考えているなら世界中でそういう話になっている。

濱島

PLCO は PSA 検診といったものについては細かい年限で区切って評価していたが、胸部レントゲンの 13 年というのは非常に恣意的に決められたエンドポイントだから、それを鵜呑みにしていいかどうか。

佐川

先生方は鵜呑みにしないでじゃなくて、ひっくり返して著者と真逆にポジティブだと言っている。

濱島

論文の評価の時点に問題があるということを提示して、それは採用できないから 6 年目の結果を評価している。

佐川

評価する権利があるのかという話。

濱島

権利じゃなくて、結果についての解釈の違い。

佐川

サブアナリシスをもってこの研究の結果だと言うのよりさらにひどい。疑義はあるが、だからと言って著者が言っていることと真逆の内容を証拠として出しても証拠にはならないのではないか。

濱島

そう考えていないから出したわけだが、6 年目の評価時点について他のレポートを見ると 6 年目で評価

すること自体特に問題はないと思う。

佐川

そう考えるならエビデンスレポートにはっきり書いてほしい。6年目を評価してそれをメインにやったと。このエビデンスレポートではそこはわからないし、症例対照研究を否定するような書き方になっている。

濱島

否定はしてない。

佐川

要するにどちらがメインかということ。

濱島

どちらがメインかと言ったら今回は PLCO がメイン。

佐川

このエビデンスレポートではそうは読めないから、それをはっきり書いてほしい。

濱島

今回、私どもは6年目の PLCO の結果を評価したということなので、症例対照研究の扱い方等についてきょうの説明も含めて詳しく説明して修正したい。

中山

千葉西総合病院の八重樫先生が挙手なさっているのでお願いしたい。

八重樫

私の意見は基本的に佐川先生と一緒に、ネガティブであるという結論をポジティブだとして後から解釈するのは問題だと思う。減弱？効果と言われているが、論文をデザインしたときにそれで検討されているならいいが、そうではないし、メインの結論はX線では肺がん死亡率は低下しないということで、他のレポートや学会でもX線は検診として効果がないと結論付けている。効果がないという根拠がおかしいと言うのなら推奨グレードを下げてIとかCにするべきで、Aにするのは全く違う。佐川先生のおっしゃる通り、ロジックの破綻があると思う。

濱島

ガイドラインの作成方法は参照して作成しているが、著者の意見と一致しないものは証拠として採用できないとは考えていなかったのが今回採用した。PLCO 自体は研究の質のいいものなので、ネガティブで採用するとなれば胸部レントゲン自体見直す方向で考えていく必要はあると思っている。

中山

チャットで「うちの検診で実施している喀痰検診をやめたいのだが良いか」というのが来ているが、きょう

はあくまでもドラフトの段階であって、厚生労働省の指針がどうなるかは全く未定なので、その点はご理解いただきたい。

佐川

喀痰のほうの話だが、D という判定はエビデンスがないということではなく、無効であるというエビデンスがあるということで、実施しているところはいますぐやめろ、任意型もやっちゃいけないということになる。おそらくこれが出たら「無効な検診をなぜ何十年もやっていたのか」ということで大騒ぎになって、おそらく誰かが責任を取ることになる。D を付けるということはそういうことで、D の概念を理解してないと思われる。

また喀痰検診をやめるとか減らすというのはあり得ない話ではないが、そうであっても有効性評価の話じゃなくて疾病構造の変化とかエキスパートの不足という方向に持っていくべき。終了する条件というのを決めておいて、まずそれを提示するというのはいいと思うが、D を付けたりしたら現場でも厚労省でも大炎上する可能性があると思う。

中山

喀痰検診でのがん発見率はX線に比べて男性は4分の1以下、女性は10分の1以下。都道府県別では西日本においてはほとんど出ていない。

佐川

地域差はある。頑張るところだとある程度見つかっているが、頑張れないところもあるだろうからそれはしょうがないものの、検診をやめたら見つからないケースも出てくる。

中山

そこは運用の話になるので、指針が変更されてからの話かと思う。

佐川

D というのは運用の話ではない。効果がないと科学的に証明されたものがD。例えば胃がん検診でピロリ菌が減っていったら、いずれ検診をやめることを考える時期も来ると思うが、それと同じで、無効だからやめるわけではない。エビデンスに基づくというのではなく、やめるための条件をきちっと出すべき。

斎藤

このDという判定だが、Iと違うことが論理的に説明できていない。もともと定義が曖昧なところはあったが、ガイドラインのグレード作成のロジックをきちんと踏まえているようには見えない。こういう面からも反論は出ると思う。

中山

ガイドライン作成委員会でもIにするかDにするかはかなり議論している。Iというのは今後何か研究が出るだろうと期待される、あるいは変更される可能性があるものであって、喀痰については海外でもやって

いないし、日本でも今後出るとは思えない。だからⅠにはできない。そうすると証拠はないと判断せざるを得ないので D としている。

斎藤

議論があったことはわかったが、40 ページの表を見ても論理的に D のカテゴリーにあてはまるかどうかあやしい。ここをきちんとエビデンスをもって論理的に説明できるようにしていただきたい。

またこのガイドライン作成の責任者が書かれてないが、誰になるのか。非常に大事なことのなのにその意識が薄い。またエビデンスプラクティスセンターという組織は回っているようだが、アドバイザリーボードという組織はどうなっているのだろうか。記載にはあるが実態があるのかどうかわからないので、そこも明記して、このガイドラインがオーソライズされたものだということをきちっと示していただきたい。

中山

修正すべきところは確実に修正したい。

佐川

先ほどの、どこにも入れるところがなかったから D に入れたという話だが、D というのは効果がないことが明らかな検診なので「我々がやった症例対照研究も全く信用が置けない」と言っているのと同じこと。Ⅰでも C でも条件を付けて、こういう状況でないとやっちゃいけないという形で対応するのならいいが、D だけはとんでもないのでこれだけはやめてほしい。

濱島

推奨グレードの判定については、アドバイザリー委員会によって証拠のレベルと 1 対 1 対応にするように修正されて、それによって大腸がん検診ガイドラインをつくっている。それ以前の子宮がん検診ガイドラインはリソース、プレファレンスを勘案してグレードを決めていたが、1 対 1 対応ではエビデンスの観点から今回では当然に D になってしまう。そのようなつくり込みになってしまったのが 1 つの要因としてある。

佐川

それは妥当だが、0.89 のものが A なのに、なんで 0.88 のものが D になるのかという話。

濱島

研究自体が違うし、規模が違うというもあるが。

佐川

両方とも大きな RCT。

濱島

もう 1 つ、この判断の背景として大事なものは、喀痰はあのプール解析以来、国内でも国外でもほとんど研究が出ていないので、あの研究が最終判断をする材料にしかない。

佐川

PLCO だってそうなるが。

濱島

そうだが、喀痰についてはあれが最終判断なので、それに基づいている。

佐川

PLCO が 0.89 で 11%の死亡率減少効果だから A で、喀痰は 0.88 で 12%の死亡率減少効果があることが想定されるけど、有意差がちょっと足りないだけで D がつくのか。

濱島

D の判断はガイドライン委員会でも正しいかどうか検討しているが、もう 1 つは 0.88 だからというところ
で、日本に外挿できるかということが大きな課題だったと思う。

佐川

私が 2003 年にケースコントロールのサブアナリシスで 1 つペーパーを書いているが、相対危険度は 0.63。
宮城だけでやったから有意差は出ていないが、外挿できるかと言ったら外挿できる。英文でパブメドに載っ
ているから引けば出てくる。

斎藤

さっきの件、アドバイザリー委員会に基づいていると言うが、がん種ごとに新たな問題に直面するからそ
の都度ブラッシュアップしてきた。大腸がんは後出しで、つくってしまったからアドバイザリー委員会に通した
形になった。今回は委員会を開いているのか。

中山

肺がんの CT の話を始めるときもエビデンスはこんな感じだとして大腸と同じ日にやっている。喀痰につい
て話す機会はなかったが、同じ日に 3 時間半ぐらいやっている。

濱島

肺がん和大腸がんに関しては特殊な事情があってほとんど並行してやっている。大腸がん和肺がんはエ
ビデンスレベルと推奨の決め方、ルールに同じものを使うと決めているので、事前に見直すことができなかつ
たというのは誤解。

斎藤

大腸和肺と一緒にやったというのなら、それは僕の勘違いで申し訳ない。ただ、これらのことをもう少し
前のタイミングで共有していただければもっと関与できたのではないかと思います。本来はそういう位置づけな
のではないだろうか。

佐川

喀痰のように検診が減ってきたものに対してどう対応すべきか、ということが新たな問題として浮上したので、そういう問題点を明確化することはできた。ただ D を付けられるのは、いままでの何十年を否定することになるので、それだけは勘弁してほしい。

中山

桜田先生がお示しいただいた末梢と肺門の割合は、2 年間追跡しても肺門は 3 例ぐらいで、末梢がかなり出てくるという理解でいいか。

桜田

割合としては 6 対 4 ぐらいのレベルで末梢のほうが多い。

中山

他に自治体の方からご意見、ご質問はないだろうか。パブリックコメントはいつまで？

細野

2 月 28 日、当部のホームページから入力できるようになっているので、きょう質問できなかったことなどがあればそちらに入力していただきたい。

中山

きょうたくさんのご意見をいただいたので、これを受けてガイドライン作成委員会で再度審議をして、改定して、また公開という形になっていくと思うが、きょういただいたご意見の多くはガイドライン作成委員会でも既に議論してきていることであって、我々が恣意的に決めたことではないということをご理解いただきたい。

それではドラフト公開フォーラムはこれでお開きにしたい。長時間ありがとうございました。

以上