

厚生労働科学研究費補助金（がん対策推進総合研究事業）
総括研究報告書

がん検診の有効性評価に関する死亡率減少につながる頑健性の高い代替指標に関する研究

研究代表者	中山 富雄	国立がん研究センター 社会と健康研究センター	検診研究部	部長
研究分担者	濱島ちさと	帝京大学 医療技術学部	教授	
	寺澤 晃彦	藤田大学 医学部	救急総合内科学	教授
	片山 貴文	兵庫県立大学 看護学部	教授	
	細野 寛代	国立がん研究センター 社会と健康研究センター	検診研究部	室長

研究要旨

がん検診の評価指標としてこれまで当該がん死亡率減少効果が用いられてきたが、前向きは無作為化比較試験では結果が得られるまで長い年数を要することから、代替指標が求められていた。本研究では海外の検診ガイドラインをレビューし、代替指標を用いている実例とその条件について検討した。検診で前がん病変の検出が容易で、前がん病変から浸潤がんへの自然歴が明らかかな子宮頸がん、大腸がんにおいて、浸潤がん罹患率が有効性評価に用いられていた。また一部のガイドラインでは両臓器において死亡率減少効果が確立している検査法との感度・特異度の比較、前がん病変・浸潤がんの検出力 (test performance) の比較で、有効性評価として行われていたが、比較する検査法を生物学的に類似した方法に限定しているものと、限定していないものが見られた。後者の場合は、感度・特異度の測定方法が様々であり絶対的な指標ではないことから、一定のルールの作成が必要である。

A. 研究目的

がん検診の評価指標としては、発見率・早期癌割合・切除率・腫瘍の大きさ・発見がんの生存率、そして当該がんの死亡率等が用いられてきた。しかし進行速度の遅いがんほど発見されやすいという特性(length-biased sampling)により、進行しても死に至らない過剰診断が発見がんに含まれることから、当該がんの死亡率以外の指標は不正確であるとされ、これまでがん検診導入の是非を検討する際には、他の指標は用いられてこなかった。がん検診の死亡率減少効果を評価する研究は、かつてはコホートや症例対照研究のような後ろ向き研究が行われてきたが、これらの研究もセレクションバイアスが混入しやすいことから、現在は無作為化比較試験が主体であり、国内でも乳がんのJ-START研究や肺がんのJEC研究が進行中である。健常者を対象としたこれらの研究は、がん死亡というイベントの頻度が少ないことから長期間かつ大規模な研究となる。研究成果が得られるのが10～15年後となるため、昨今の医療技術の開発速度の速さとは相容れず、有効性が確認された時点で

は、その技術が新しい技術に押されて陳腐化しているのではないかという批判がこれまでもあった。研究期間の短縮のための代替指標の必要性についての議論が期待されている。代替指標として浸潤がん罹患率、前がん病変の罹患率、転移がん罹患率、stage shift、感度・特異度などが候補に上がるが、これらを評価指標に含めた研究はこれまでも散発的に報告されていたものの死亡率減少効果に確実に紐づくかは定かではない。また測定の方法によってバイアスが混入するかも検証されていない。本研究では、がん検診の有効性を評価する指標の扱いや代替指標及び具体的な研究手法について、国際学会や海外のガイドラインの動向、文献等から情報を収集し、死亡率減少効果に確実に紐づく代替指標とその検証方法および代替指標が使える条件等について整理し、「がん検診の有効性評価のための指標と研究手法に関する手引」としてまとめることを目的とする。

B. 研究方法

本研究においては、がん検診の死亡率減少効果以外の代替指標について、諸外国のガイドラインが推奨の判断材料として検討対象としているか？またそれぞれの代替指標にバイアスが混入するかの、死亡率減少効果との間で結果の乖離がないかを文献レビューと、ガイドライン作成組織への直接調査により把握し、具体的な代替指標とそれを適応する条件を検討した。

具体的には以下のスケジュールで検討を行なった。本年度は初年度にあたり、諸外国のがん検診あるいはがんに関する診断法に対してのガイドラインを収集し、評価対象とされているエンドポイントを確認した。対象として常設の予防ガイドライン作成組織である Cochrane 共同計画、米国の US Preventive Service Task Force、American Cancer Society、National Comprehensive Cancer Network、英国の UK National Screening Committee、カナダの Canadian Task Force とした。また IARC、WEO のガイドラインについても検討した。これらの対象となるガイドラインで、死亡率減少効果以外のエンドポイントを評価対象として検討されたものがないかを検索した。

代替指標として候補に上がるものは、浸潤がん罹患率、前がん病変の罹患率、転移がん罹患率、stage shift、感度・特異度などが考えられることから、これらを評価指標とした研究が、ガイドラインで推奨の判断材料として考慮されているかどうかを確認した。すでに子宮頸がん検診においては、前がん病変からの浸潤がんへの自然史が明らかになっていることから浸潤がん罹患率を評価指標とした研究が以前から行われていた。また大腸がんに関しては、便潜血検査化学法については、死亡率減少効果が複数の無作為化比較試験により示されているものの、それ以外の検診手法については、化学法との比較で、推奨が検討されている。その際の比較する指標・考え方をレビューした。

（倫理面への配慮）

本研究は既出の文献、ガイドラインの検討であり、倫理面の問題は発生しない。

C. 研究結果

国際的なガイドラインとして、大腸・子宮頸・乳房について USPSTF、ACS、NCCN のガイドラインを入手し、レビューを行った。自然歴が確立し、既存の検診手法の死亡率減少効果が確定している

大腸・子宮頸がんに関しては、浸潤がん罹患率減少効果を用いたり、死亡率減少効果が確定している検診手法との精度比較で推奨を決めていた。そこで用いられていたのは、IARC が 2007 年に公表した European guideline for quality assurance of cervical cancer screening (second edition) である。ここでは表 1 に示すように、死亡率減少効果を優先度を 1 番に置き、2 番目に浸潤がん（ここでは stage IB）罹患率減少効果として、評価指標の順番を設定している。2018 年に公表された USPSTF ではこの考え方を活用し、浸潤がん罹患率減少効果をエンドポイントとした HPV 検査を用いた子宮頸がん検診の評価を行っている。

大腸では、便潜血検査化学法と S 状結腸鏡については、死亡率減少効果を指標とした無作為化比較試験が行われていたが、便潜血免疫法、注腸検査、全大腸内視鏡検査、CT コロノグラフィなどは、検討したガイドラインにおいて、便潜血検査化学法との精度比較を基に推奨を決めていた。USPSTF(2016)では、systematic review で死亡率減少効果を評価していたが、対象となったのは便潜血免疫法と S 状結腸鏡にとどまり、他の手法については研究自体がなかった。主として test performance が評価されていた。ここでの test performance は、前がん病変である advanced adenoma と大腸がん腺腫の検出力を既存の死亡率減少効果が確立している便潜血検査化学法と比較することであった。大腸がん検診ガイドライン自体には、代替指標の使い方について明記はされていないものの、死亡率に関する評価研究自体がなかった CTC、全大腸内視鏡検査が推奨されており、考察には「どの検査が優れているのかという評価は行わなかった。米国では、対象となる成人の約 3 分の 1 が大腸がん検診を受けたことがない」とされており、大腸がん検診の方法を選択できるようにすることで、検診の受診率が向上する可能性がある。そのため、検診項目は優先順位やランク付けされたものではなく、むしろ、検診を受ける人の総数を最大化することが目的であり、そのことが大腸がんの死亡数の減少に最大の効果をもたらすと考えられる、と書かれていた。ACS の 2018 年に公表された Colorectal cancer screening for adults average-risk adults では、USPTSF に準じたという但し書きを示した上で、罹患率・死亡率の低下、test performance、偶発症を検討しているが、罹患率・死亡率に関する研究手法として、モデル研究を許容していた。

WEO (World Endoscopy Organization) でも、すでに評価の確立した方法と新技術の感度・特異度の比較を評価の第一段階に位置付けていたが、評

価はそこでとどまらず、最終アウトカムによる評価を継続して行うことが望まれていた。またここで比較していた新技術は評価の確立した方法と生物学的機序が類似したものに限定され、全く異なる機序による検査方法は対象外とされていた。IARCもWEOの考え方を踏襲し、生物学的に類似の検査は感度・特異度により有効性を検証可能という見解を示していた。また中間期がんの評価の重要性を示していた。

乳がん検診においては、いずれのガイドラインも代替指標を評価していなかった。

D. 考察

今回検討した各国のガイドラインでも死亡率減少効果を最終アウトカムとして、評価はしていたが、子宮頸がん・大腸がんという前がん病変が検診で容易に発見され、かつ前がん病変の治療が浸潤がんの罹患の減少・死亡率減少につながるという臓器において、代替指標が評価されていた。子宮頸がんでは浸潤がん罹患率が中間アウトカムとして評価することが広く行われていたが、これは上記の自然歴が明らかなことと、先進国での死亡率減少が顕著で死亡率減少効果を評価する研究が現実的ではないことにも影響されている。ただし比較する対照群との間で、他のリスク要因の分布が異なると、その影響を被るため、無作為化比較試験による評価が必須であり、観察研究の場合エビデンスレベルは低くなることには注意が必要である。大腸がんの場合、感度・特異度を中間アウトカムとして評価することが行われているが、後発であるWEO、IARCのように評価する新技術を、有効性が確立している既存の技術と生物学的に類似した技術に限定して使用しているガイドラインと、USPSTFのように限定せず、死亡率減少効果が確立している便潜血化学法と感度・特異度を比較することで有効性の指標としているガイドラインがあった。後者がこのような評価を最初に用いたのは、2016年であり、その前の2008年版では、死亡率減少効果のエビデンスのないS状結腸鏡や全大腸内視鏡を数理統計モデルで評価していたが、2016年版では同様に数理統計モデルで利益・不利益の評価もしていたが、感度・特異度の評価のウエイトが大きかった。議論の根拠は示されていないものの、ガイドライン自体には、「米国での大腸がん検診受診率が低いので、どの検診手法が優越であるのかの順位付けは行わなかった」と明記されており、受診率向上対策のため選択肢を広げるため判断基準を広げた可能性が推察された。感度・特異度を代替指標として用いた場合の問題

点は、その値が絶対的なものではないということである。感度・特異度の測定方法もさまざまであり、大腸がん検診の評価で通常用いられる同時法の他、がん登録と照合する追跡法がある。同時法は同時に行われるgold standardの精度との相対的な値であり、gold standardの精度を100%とみなしたものであるため、追跡法の値より大きな値となる。よって両者を比較することは妥当ではない。無作為化比較試験の形式で既存の死亡率減少効果が確立した手法と新技術を比較する研究が行われているが、これも同時法と同様の相対的な値となる。また診断に至るまでの精密検査の精度も問題となる。一般的には、検診手法単体の感度はtest感度、精密検査を含めた感度はepisode感度と呼ばれる。また追跡法・同時法はいずれも発見がんを分子として感度が計算されるが、発見がんには必ず過剰診断が含まれる。このため過剰診断を除外した罹患率法(incidence method)も最近では推奨されているが、まだ報告は乏しい。更に新技術の中には、患者群と健常者群を別個に集めて1:1などの人為的な比率で擬似的に作成した集団で検討するtwo-gate design(case-control design)のみが報告されている場合が多い。本手法は研究費の軽減のため開発段階で広く行われているが、患者群と健常者群の比率を変化させることで、感度も特異度も100%に近くなることが以前から指摘されており、診断精度の評価ツールであるQUADAS-II(Quality Assessment tool for Diagnostic Accuracy Studies-II)においても、two-gate designによる精度評価は用いるべきではないとされている。このように代替指標として感度・特異度を用いる場合は、その測定方法が様々であることから、一定のルールが必要となる。またtest performanceは、発見率の比較である。incidence法以外の感度およびtest performanceはいずれも発見病変を分子としているため、過剰診断を含み、過大評価につながりやすい。このためIARCは中間期がんの評価の重要性を示している。次年度はこれらの代替指標の活用に関するルールを設定を課題とする。

E. 結論

前がん病変の検出が容易かつ自然歴が明らかな子宮頸がん・大腸がんに関しては、中間アウトカムとして浸潤がん罹患率と感度・特異度を代替指標として有効性評価に含めている国際的なガイドラインを認めた。ただし感度・特異度に関しては、測定方法が様々であり、一定のルールが必要である。

F. 健康危険情報
特になし

G. 研究発表

1. 論文発表

1. Aoe J, Ito Y, Fukui K, Nakayama M, Morishima T, Miyashiro I, Sobue T, Nakayama T. Long-term trends in sex difference in bladder cancer survival 1975-2009: A population-based study in Osaka, Japan. *Cancer Med* 2020 ; 9(19) : 7330-7340. Doi: 10.1002/cam3.3382
 2. Yagi A, Ueda Y, Matsuda T, Ikeda S, Miyatake T, Nakagawa S, Hirai K, Nakayama T, Miyagi E, Enomoto T, Kimura T. Japanese mothers' intention to HPV vaccinate their daughters: How has it changed over time because of the prolonged suspension of the governmental recommendation? *Vaccine* 2020; 8(3): 502; doi: 10.3390/vaccines8030502.
 3. Nakagiri T, Nakayama T, Tokunaga T, Takenaka A, Kunoh H, Ishida H, Tomita Y, Nakatsuka SI, Nakamura H, Okami J, Higashiyama M. Novel Imprint Cytological Classification for Small Pulmonary Adenocarcinoma Using Surgical Specimens: Comparison with the 8th Lung Cancer Staging System and Histopathological Classification. *J Cancer*. 2020 Feb 21;11(10):2845-2851. doi: 10.7150/jca.35027. eCollection 2020. PMID: 32226502
 4. Nakagiri T, Nakayama T, Tokunaga T, Takenaka A, Kunoh H, Ishida H, Tomita Y, Nakatsuka S, Nakamura H, Okami J, Higashiyama M. Intraoperative Diagnosis and Surgical Procedure with Imprint Cytology for Small Pulmonary Adenocarcinoma *J Cancer* 2020; 11(10):2724-2729. doi:10.7150/jca.35026
 5. Kono K, Morisada T, Saika K, Saitoh-Aoki E, Miyagi E, Ito K, Takahashi H, Nakayama T, Saito H, and Aoki D. The first-round results of a population-based cohort study of HPV testing in Japanese cervical cancer screening: baseline characteristics, screening results, and referral rate. *J Gynecol Oncol*. (Impact factor=3.304) 2021;32:e29. Published online Jan 26, 2021.
 6. Taniguchi M, Ueda Y, Yagi A, Miyoshi A, Tanaka Y, Minekawa R, Endo M, Tomimatsu T, Hirai K, Nakayama T, Kimura T. Disparity of Cervical Cancer Risk in Young Japanese Women: Bipolarized Status of HPV Vaccination and Cancer Screening. *Vaccines (Basel)*. 2021 Mar 19;9(3):280. doi: 10.3390/vaccines9030280. PMID: 33808630
 7. Chisato Hamashima. The burden of gastric cancer. *Ann Transl Med*. 2020 Jun;8(12): 734. doi: 10.21037/atm.2020.03.166. PMID: 32647659 PMCID: PMC7333126 DOI: 10.21037/atm.2020.03.166
 8. Hamashima C, Yoshimura K, Fukao A. A study protocol for expanding the screening interval of endoscopic screening for gastric cancer based on individual risks: prospective cohort study of gastric cancer screening. *Ann Transl Med*. 2020 Dec;8(23):1604. doi: 10.21037/atm-20-5949. PMID: 33437803
 9. Hirai K, Ishikawa Y, Fukuyoshi J, Yonekura A, Harada K, Shibuya D, Yamamoto S, Mizota Y, Hamashima C, Saito H. Correction to: Tailored message interventions versus typical messages for increasing participation in colorectal cancer screening among a non-adherent population: A randomized controlled trial. *BMC Public Health*. 2020 Nov 16;20(1):1730. doi: 10.1186/s12889-020-09823-x. PMID: 33198697
2. 学会発表
1. 中山富雄、HPVを用いた子宮頸がん検診のエビデンス。第72回日本産婦人科学会総会 子宮頸がん検診事業を考えるワークショップ、2020.4.15
 2. 中山富雄、地域や施設での検診精度を保つために何をすればいいのか? JDDW2020 要望講演、2020.11.5
 3. 中山富雄、わが国の肺がん検診の現状。第61回日本肺癌学会総会、WS-2 Aiming to Reduce

Lung Cancer Mortality- Lung Cancer Screening; What's Next?, 2020.11.12、岡山市

4. 中山富雄、肺がん検診における精度管理の重要性、第61回日本肺癌学会総会、肺がん検診セミナー 基調講演、2020.11.14、岡山市
5. 寺澤 晃彦. HPV 検査を用いた検診のエビデンス. ガイドラインセッション「有効性評価に基づく子宮頸がん検診ガイドライン更新版と今後の課題」. 第 29 回日本婦人科がん検診学会総会・学術講演会. 2021/2/20
6. 細野 覚代. 細胞診を用いた検診の不適正割合・対象年齢・検診間隔について. ガイドラインセッション. 第 29 回日本婦人科がん検診学会学術総会 2021.02.20. 新潟オンライン
7. 細野 覚代. 大学病院勤務者におけるB型肝炎ワクチン接種効果の検討. 一般口演. 第79回日本公衆衛生学会総会 2020.10.20. 京都 オンライン

H. 知的財産権の出願・登録状況
(予定を含む。)

1. 特許取得
なし
2. 実用新案登録
なし
3. その他
なし

表 1. 子宮頸がん検診のアウトカムと証拠のレベル

証拠のレベル	アウトカム
ランク 1	子宮頸がん死亡の減少
ランク 2	子宮頸がん罹患(病期 1B+の発症)の減少
ランク 3	子宮頸がん罹患(微小浸潤がんを含む)の減少
ランク 4	CIN3+病変の減少
ランク 5	CIN3+病変(または CIN2+病変)の検出増加 ①累積 CIN3+病変検出増加 ②CIN2 病変検出増加確認後の検診で CIN3+病変の減少
ランク 6	スクリーニング陽性数の増加(陽性適中率の増加または不変を必要とするが、わずかの減少まで許容する)

(European guideline for quality assurance of cervical cancer screening (second edition) by IARC より引用)

表 2. USPSTF 大腸がん検診ガイドラインの変遷

発表年	推奨の内容と評価の変遷
2002	死亡率減少効果が確立していない全大腸内視鏡、注腸検査は証拠不十分として推奨せず
2008	死亡率減少効果が確立していない全大腸内視鏡、便潜血免疫法、便潜血とS状結腸鏡併用法を推奨。 数理モデルにより、利益・不利益バランスを評価
2016	死亡率減少効果が確立していない全大腸内視鏡、便潜血免疫法、便潜血とS状結腸鏡併用法、便DNA検査、CTコロノグラフィを推奨。 数理統計モデルにより、利益・不利益バランスを評価。感度・特異度の評価を充実させた。

表3 大腸がんの検査法比較の評価 (Young et al. Recommendations for a Step-Wise Comparative Approach to the Evaluation of New Screening Tests for Colorectal Cancer. Cancer 2016;)

評価相		第一目的	第二目的	対象集団
第1相	検診前:がんと健常者の判別が可能かの後ろ向き評価	がんの発見能 1.1 精度の主要尺度としての真陽性率と真陰性率の評価	1.2 収集法の過程の確立 1.3 品質管理の最適化 1.4 検査の指標の調整	無症状者を含み救命可能な状況のものが大多数を占めるがん患者とがんと持たないものを対照とする。理想的には診断過程ですべての結果を確認した上でペア検査が望ましい。
第2相	前向き臨床試験:一連の腫瘍性病変の発見	顕在化する前の初期腫瘍の発見能 2.1 発がん過程に沿った病期の腫瘍、特に advanced adenoma を含む前臨床期病変の発見精度 2.2 検査の最終的なフォーマットの確定(標本と指標) 検査結果登録の最小限の装備	2.3 より信頼性の高い精度特性の評価 2.4 診断精度に影響する因子の情報 2.5 標本数と閾値 2.6 信頼性のある検査結果の登録 2.7 発見できないサブグループの有無の明確化	症例は特に早期癌や advanced adenoma を含んでいて、症状の有無の情報があること。対照は腫瘍がなく、良性疾患を有するものの場合、検査結果に影響が出るかを確認する必要がある。診断過程の前に行われる検査であり、理想的にはペア検査で、比較する検査と同時に報告される。
第3相	初回検診の評価; 第一回目検診	検診発見病変の性質; 偽陽性率、受容性 3.1 検診受診者で検査の精度特性、偽陽性を含む発見に関連した指標の収集 3.2 受容性の評価	3.3 検診発見病変の性質と頻度 3.4 実行可能性 3.5 診断過程を含む費用の事前評価	初回検診で典型的な検診の環境での検査; 別々のコホートで新しい検査または比較対象(通常の治療の場合もある)を実施し、受診勧奨からアウトカムまでを追跡; 検査陽性者のみが全大腸内視鏡を必要とする(全大腸内視鏡との直接比較の場合を除く); 最初は小規模の単純な方法の研究で始めて、発見率を求める大規模研究に移行する; intention-to-screen 分析が必要。
第4相	複数回検診の評価	腫瘍病変の減少の評価と偶発症 4.1 がん死亡率減少効果の評価、あるいはモデルでの評価	4.2 広義の利益 4.3 正確な費用 4.4 2回目以降の検診への参加 4.5 診断までのフォローアップ遵守率 4.6 発見した病変の治療可能性 4.7 検診間隔 4.8 偽陰性率	検診プログラムとして実施される可能性のあるような集団から無作為に抽出; 研究デザインとしてヒストリカルコントロールか、別の検診手法との無作為比較試験; intention-to-screen 分析が必要

			<p>4.9 繰り返し検診でのプログラム発見率</p> <p>4.10 全検診でのフォローアップ率</p> <p>4.11 病変発見のNNS (number needed to screen)</p> <p>4.12 予期せぬ偶発症</p>	
--	--	--	--------------------------------------------------------------------------------------------------------------------------	--