

有効性評価に基づく がん検診ガイドライン作成手順



平成16年度 厚生労働省がん研究助成金
「がん検診の適切な方法とその評価法の確立に関する研究」班

有効性評価に基づく がん検診ガイドライン作成手順

2005 年 3 月 24 日

平成 16 年度 厚生労働省がん研究助成金

「がん検診の適切な方法とその評価法の確立に関する研究」班

主任研究者 祖父江友孝

目次

研究班構成	1
要旨	3
Ⅰ. はじめに	4
Ⅱ. 本ガイドラインにおける基本的な考え方	6
Ⅲ. ガイドライン作成の概要	8
1. ガイドライン作成の手順	8
2. 検診の対象	8
3. 有効性評価の指標	8
Ⅳ. Analytic Framework の作成	10
1. Analytic Framework の作成	10
2. 直接的証拠と間接的証拠	10
Ⅴ. 文献検索	11
1. 検索方法	11
2. 対象文献の選択	11
Ⅵ. 個別研究評価	13
Ⅶ. 証拠のまとめ	14
1. 証拠のレベル	14
2. 不利益の評価	15
Ⅷ. 推奨への翻訳	16
Ⅸ. ガイドライン・ドラフトの作成	18
1. ガイドライン・ドラフトの作成	18
2. 経済評価及び検診実施の障壁関連の個別研究評価	18
Ⅹ. 外部評価	20
1. 外部評価の手順	20
2. 外部評価チェックリスト	20
3. 公開フォーラムの開催	21
XⅠ. ガイドラインの公開と再評価	22
XⅡ. 今後の課題	23
文献	26

図表

図 1 : ガイドライン作成過程.....	31
図 2 : がん検診の Analytic Framework と検討課題.....	32
表 1 : 久道班第 3 版報告書の評価判定	33
表 2 : 久道班第 3 版報告書の根拠の質	34
表 3 : 対策型検診と任意型検診の比較	35
表 4 : 証拠のレベル.....	36
表 5 : 大腸がん検診の不利益.....	37
表 6 : 推奨のレベル.....	38
添付書類	
添付書類 1 : チェック・リスト解説.....	39
添付書類 2 : 研究方法別チェック・リスト	53
添付書類 3 : 外部評価チェック・リスト（胃がん検診の例）	61
添付書類 4 : 用語解説	77

研究班構成

平成 15-16 年度 厚生労働省がん研究助成金 (15-3)

「がん検診の適切な方法とその評価法の確立に関する研究」班

主任研究者

祖父江友孝 国立がんセンター がん予防・検診研究センター 情報研究部 部長

分担研究者

深尾 彰 山形大学医学部環境病態統御学講座公衆衛生・予防医学分野 教授
辻 一郎 東北大学大学院医学系研究科医科学専攻社会医学講座公衆衛生学分野 教授
大貫幸二 岩手県立中央病院 乳腺外科 科長
佐川元保 金沢医科大学呼吸器外科 助教授
青木大輔 慶應義塾大学医学部産婦人科 教授
齋藤 博 国立がんセンター がん予防・検診研究センター 検診技術開発部 部長
中山富雄 大阪府立成人病センター調査部疫学課 参事
渡邊能行 京都府立医科大学大学院医学研究科地域保健医療疫学 教授
田中純子 広島大学大学院医歯薬学総合研究科展開医科学専攻病態情報医科学講座
疫学・疾病制御学 助教授

研究協力者

坪野吉孝 東北大学大学院法学研究科公共法政策 教授
島田剛延 宮城県対がん協会がん検診センター 副所長
本荘 哲 栃木県立がんセンター研究所疫学研究室 室長
中山健夫 京都大学大学院医学研究社会健康医学系専攻健康情報学分野 助教授
松田一夫 財団法人 福井県健康管理協会県民健康センター 所長
西田 博 松下健康管理センター 副所長

顧問

久道 茂 宮城県病院管理者
大島 明 大阪府立成人病センター調査部 部長

事務局

濱島ちさと 国立がんセンター がん予防・検診研究センター情報研究部 室長
吉見逸郎 国立がんセンター がん予防・検診研究センター情報研究部 研究員

本ガイドライン作成は、平成 15-16 年度 厚生労働省がん研究助成金 (15-3) の助成を受けた。研究班のすべての構成員は、大腸がん検診に関与する医療器械・医薬品・検査試薬等の会社から、研究費等の支援は受けていないことを明記する。

平成 15-16 年度 厚生労働省がん研究助成金「がん検診の適切な方法とその評価法の確立
に関する研究」班 ガイドライン作成手順検討委員会

祖父江友孝	国立がんセンター	がん予防・検診研究センター	情報研究部
濱島ちさと	国立がんセンター	がん予防・検診研究センター	情報研究部
齋藤 博	国立がんセンター	がん予防・検診研究センター	検診技術開発部
中山健夫	京都大学大学院医学研究社会健康医学系	専攻健康情報学分野	
中山富雄	大阪府立成人病センター	調査部疫学課	

要旨 背景

わが国では、平成10年から過去3回に亘り、がん検診の有効性評価が行われた。大半の市区町村の老人保健事業では有効と判定されたがん検診が行われているが、一部の市区町村、職域、人間ドック等では、保留あるいは無効と判定されたがん検診を含め多様な検診が行われている。一方、近年、欧米を中心にガイドラインの標準化が進められつつある。

目的

今後、ガイドライン更新を行っていくために、その方法を明確にし、定式化する。

ガイドラインの概要

無症状の一般的な健常者を対象としたがん検診による死亡率減少効果を明らかにするため、最新の知見も含めた対象文献の系統的総括を行う。各検診方法の死亡率減少効果と不利益に関する科学的根拠を示し、わが国において集団を対象とした対策型検診及び個人を対象とした任意型検診として実施の可否を推奨として総括する。さらに、がん検診に関与するすべての人々への有効性評価に関するガイドラインの情報を提供する。

死亡率減少効果の証拠

死亡率減少効果について、直接的証拠と間接的証拠の両者を採用する。ただし、間接的証拠は、単独ではなく、直接的証拠のある検診方法との比較検討が可能な場合にのみ、採用する。ここでいう直接的証拠とは、無作為化比較対照試験により死亡率減少効果を証明した研究である。一方、間接的証拠とは、個々の研究だけでは検討対象となるがん検診による死亡率減少効果を証明することが困難なことから、複数の研究の集積により死亡率減少効果が示唆されるものである。

ガイドライン作成手順

がん検診ガイドラインの作成は、対象となるがん検診を選定し、Analytic Frameworkに基づき、検討課題を明らかにする。評価のための文献収集を行い、チェック・リストを用いて、個別研究の評価を行う。個別研究は検診方法別にエビデンス・テーブルとして統括され、証拠のレベルと不利益に基づき、推奨が決定される。ガイドライン・ドラフトは、外部評価、公開フォーラムでの討議を経て、ガイドラインとして公開される。ガイドラインは、その理解度や普及度などについてアンケート調査で定期的にモニターする。調査結果を参考に、新たな検診方法の評価を加え、5年以内にガイドラインが更新される。

推奨の過程

対象となる検診の直接的証拠と間接的証拠を検討し、8段階に分類される証拠のレベルを決定する。対象となるがん検診の特性を考慮し、各検診方法別の不利益についての比較表に基づき、委員会内で検討する。証拠のレベルと不利益を考慮し推奨レベルが決定される。推奨は、集団を対象とした対策型検診と個人を対象とした任意型検診として実施の可否を総括する。すなわち、死亡率減少効果を認め、集団及び個人に実施を勧めるA及びB、死亡率減少効果を認めるが、不利益が無視できないことから、集団には実施を勧めないC、死亡率減少効果を認めないことから、実施を勧めないD、証拠不十分で判定保留のIである。

I. はじめに

がん検診によりがん死亡を減少させるためには、有効性の確立した検診を正しく実施する必要がある。がん検診の有効性評価については、平成10年3月の厚生省老人保健推進費補助金 老人保健福祉に関する調査研究等事業「がん検診の有効性評価に関する研究班」報告書（主任研究者 久道茂）¹⁾をはじめとして、平成11年3月、平成13年3月と過去3回にわたる評価が行われている。その目的は、がん検診を公的施策として行う自治体に判断材料を提供することであった。これらの報告書は厚生労働省のがん検診検討委員会による審査を経て、全国の市町村に配布され、老人保健事業によるがん検診に利用されている。

現在、大半の市区町村における老人保健事業では、平成13年3月に公表された平成12年度厚生労働省老人保健事業推進費等補助金 がん検診の適正化に関する調査研究事業「新たながん検診手法の有効性評価」報告書（主任研究者 久道茂、以下、久道班報告書第3版）²⁾により有効と判定されたがん検診が行われている。しかし、一部の市区町村、職域、人間ドックなどでは、保留あるいは無効と判定されたがん検診を含め、多様な検診が行われている。

久道班報告書第3版における評価

平成13年3月に公表された久道班報告書第3版においては、老人保健事業によるがん検診の他、新たに行われている検診方法を含め、系統的総括を行い、根拠の質に基づき、評価判定を行った²⁾。総合評価は、検査精度・検診発見がんと臨床診断がんの比較・死亡率減少効果・経済効果・不利益に関する個々の研究をまとめ、検診による死亡率減少効果の有無を判断するものである（表1）。さらに、死亡率減少効果の有無の根拠にも各々2段階に評価を行い、判断不能なものは保留（Ⅱ群）としている。

根拠の質は、評価判定の根拠となった研究方法を示している（表2）。研究方法の妥当性は、死亡率減少効果を示す無作為化比較対照試験の評価が最も高く、死亡率減少効果に関する検討が行われていない場合には、5と判定されている。評価方法は主として死亡率減少効果を示す研究の研究方法に基づくものであり、他の要因については傍証的な位置づけとなっている。

有効性評価の動向

2000年にLancetに公表された乳がん検診に関するGøtsche PCらの再評価は、無作為化比較対照試験による評価に疑義を投げかけるものであった³⁾。これまでがん検診のガイドラインにおける推奨は、主として研究方法に立脚した評価であり、研究の質や不利益は次善の基準にすぎなかった。しかし、今後、研究方法のみを重視した評価ではなく、研究の質や利益と不利益のバランスを考慮した評価が求められるようになった。このため、US Preventive Services Task Force（以下、USPSTF）を始めとする臨床ガイドラインやデータベースなどで、乳がん検診の再評価が行われた。相前後して、ガイドラインの推奨

基準にも利益と不利益のバランスを考慮されることが基本となってきた。

2001年にはECを中心としてAGREE (Appraisal of Guideline for Research and Evaluation)⁴⁻⁶⁾が発足し、引き続き、GIN (Guidelines International Network)⁷⁾、COGS (Conference on Guideline Standardization)⁸⁾⁹⁾、GRADE (Grade of Recommendation, Assessment, Development, and Evaluation) Working Group¹⁰⁾により、臨床ガイドラインの標準化や推奨基準の見直しが提唱され、ガイドライン作成についての国際共同研究が進められるようになった。

わが国において、これまで行われてきたがん検診の評価は、研究班を主体としていたため、必ずしも定期的な評価の見直しや更新が予定されていたわけではない。また、久道班報告書第3版²⁾はUSPSTF第2版¹¹⁾の手順を参考にし、評価判定が行われていたが、文献検索の方法、系統的総括の過程や推奨ルールは明確化されておらず、ガイドラインとしての体裁は十分とはいえなかった。諸外国では、公的常設機関により臨床ガイドラインの作成・更新が行われており、その詳細な手順も公表されている。今後、わが国においても、諸外国の動向を見据え、最新の知見に基づき、有効性評価の継続的な更新を行い、政策決定の基礎資料を提供する必要がある。

Ⅱ. 本ガイドラインにおける基本的な考え方

がん検診による早期発見・早期治療による死亡率減少効果は、わが国の国民全体の健康状態の改善に大きな利益をもたらすものである。しかし、がん検診は診療とは異なり、健常者を対象としていることから、利益ばかりでなく、様々な不利益がある。このため、予防対策としてがん検診を行う場合には、利益と不利益のバランスを考慮し、意思決定を行わなければならない。個人の任意でがん検診を受診する場合も、同様に利益と不利益について考慮しなくてはならない。

がん検診による死亡率減少効果を明らかにするため、最新の知見も含めた関連文献の系統的総括を行う。その結果に基づき、各検診方法の死亡率減少効果と不利益に関する科学的根拠を明確にし、わが国において集団及び個人を対象としたがん検診としての実施の可否について推奨として総括する。集団を対象とする検診は公共的施策としておこなわれることから対策型検診 (Organized Screening)、個人を対象とした検診は任意型検診 (Opportunistic Screening) と定義する。この概念は、欧米における組織化された Organized Screening や、かかりつけ医主体の Opportunistic Screening とは、厳密な意味では相違する点があるが、がん検診を推進する上で意図するところはほぼ同一の概念と考えられる。両者の特性については、Miles らが指摘しているように、検診の目的や方法、利益・不利益の考え方が異なっている¹²⁾ (表 3)。

集団を対象としたがん検診とは、集団の死亡率減少を目的として実施するものを示し、公共的な予防対策として行われる。このため、偶発症や受診者の負担などの不利益を最小とするのが基本条件となる。集団の死亡率減少を目的とすることから、検査方法として感度のみが優先されるのではなく、特異度もあわせて検討する必要がある。対策型検診は、市町村が行う老人保健事業による集団検診・個別検診や職域の法定健診に付加して行われる検診が該当する。北欧や英国では、対象者名簿を有し、系統的な勧奨や追跡調査を行う、組織化されたがん検診 Organized Screening が行われている。しかし、わが国における対策型検診は必ずしも同義ではなく、市町村格差もあり、未整備の状況にある。一方、個人を対象とした任意型検診とは、個人の任意により受診するがん検診を意味する Opportunistic Screening に相当する。任意型検診は、個人の死亡リスクの減少を目的としているが、対策型検診と同様に死亡率減少効果を示す根拠が必要である。利益が不利益を優先する場合が多いため、感度が高い検査法が選択されるが、不利益も無視することはできない。対象となる個人は通常の診療の範囲外となる健常者であり、具体的には、検診機関や医療機関などで行われている総合健診や人間ドックなどに含まれているがん検診が該当する。

本ガイドラインは、対策型検診・任意型検診にかかわらずがん検診に関連するすべての人々への情報提供を目的としている。すなわち、がん検診の計画立案や実施マネジメントに関与する保健医療の行政職、医師、保健師、看護師などの保健医療職、事務担当者、

さらに、がん検診の受診者を対象としている。このため、がん検診の提供を検討するすべての医療機関において活用されることを期待すると共に、がん検診受診者への適切な情報提供を目指すものである。

本研究班の提示する推奨は、あくまで死亡率減少効果と不利益に関する科学的証拠に基づいた判断である。推奨すると判断したがん検診についても、集団を対象とした対策型検診として実際に導入する場合、対象集団での罹患率、経済性、利用可能な医療資源、他の健康施策との優先度など、他の多くの要因を考慮すべきである。本研究班で推奨すると判断した検診を実際には導入しないことが合理的である場合はあるが、推奨しないと判断した検診を導入することは、多くの場合、合理的ではない点に留意すべきである。

Ⅲ. ガイドライン作成の概要

1. ガイドライン作成の手順

ガイドラインの作成にあたり、主として USPSTF 第3版¹³⁾、米国 CDC (U.S. Centers for Disease Control and Prevention) による Guide to Community Preventive Services¹⁴⁾、英国イングランド・ウェールズにおける NICE (National Institute for Clinical Excellence)¹⁵⁾、英国スコットランドの SIGN (Scottish Intercollegiate Guidelines Network)¹⁶⁾ における臨床ガイドラインの作成手順及び臨床ガイドラインのチェック・リストである AGREE instrument⁴⁻⁶⁾ を参考にし、わが国のがん検診ガイドライン作成の基本となる手順をまとめた。この作成手順を基本として、わが国におけるがん検診について定式化された方法により評価を行っていく予定である。

がん検診ガイドライン作成は、作成委員会編成に始まる。委員会のメンバーは、対象となるがん検診の専門家・実務担当者や、当該がんの診断・治療に従事する臨床医をはじめとする医療従事者、疫学者、医療経済学者など幅広い分野から選出される。ただし、ガイドライン作成の中心的な機能を担う常設機関が存在しない現段階では、その役割を本研究班が担う。検討対象となる文献の抄録チェックや個別研究の評価などガイドラインに関わる作業は、分担研究者、研究協力者が共同で行い、事務局にて統括する。

図1に示すように、がん検診ガイドラインの作成は、対象となるがん検診を選定し、その Analytic Framework に基づき、検討課題を明らかにする。評価のための文献収集を行い、抄録のチェックを経て、チェック・リストに従い、個別研究の評価を行う。個別研究は検診方法別にエビデンス・テーブルとして統括され、死亡率減少効果を示す証拠のレベルと不利益に基づき、推奨が決定される。この時点でまとめられたガイドライン・ドラフトは、外部評価や公開フォーラムにおける検討に基づき、追加・修正を行い、ガイドラインとして公開される。公開されたガイドラインは、その理解度や普及度などについてアンケート調査や官庁統計など各種の公開されている情報を用いて定期的にモニターする。調査結果を参考に、さらに新たな検診方法の評価を加え、ガイドラインが更新される。対象となるガイドラインの作成開始から、公表までの期間は12~18か月を標準とする。

2. 検診の対象

基本的には、検診の対象は、無症状の一般的な健常者である。高リスクな集団を対象とする場合には、その旨を明記する。対象となるがん検診および検診方法の詳細は、ガイドライン作成開始時に、ガイドライン作成委員会の合議で決定する。

3. 有効性評価の指標

がん検診の有効性を評価するためには、適切な指標を設定する必要がある。がん検診の

評価指標は、対象となるがんの死亡率である。発見率や生存率などの指標が、がん検診の有効性を直接示す指標とはならない。発見率は、対象集団のがんの罹患や滞在時間に影響を受ける。また、生存率は、検診固有のバイアスであるリードタイム・バイアスやレンジス・バイアスの影響を排除できないためである。両者共に単独では有効性を示す指標にはならないが、検診精度など他の指標との組み合わせにより、間接的証拠の一部に用いられる場合もある。本ガイドラインでは、がん検診の死亡率減少効果を示す直接的証拠と、検診の精度や生存率などの複数の研究を積み重ねることにより得られる間接的証拠の両者を、有効性評価の根拠として採用している。

IV. Analytic Framework の作成

本ガイドラインでは、死亡率減少効果について、直接的証拠と間接的証拠の両者を採用する。

1. Analytic Framework の作成

各がん検診の Analytic Framework (図 2、以下、AF) を作成する。AF とは、単に検診の流れを整理したものではない。検査や治療の結果を評価するために、検診、精密検査、治療の段階において、各段階における評価指標 (検診による中間結果) を明確にし、最終的な結果である死亡率減少にどのように結びついていくかという、一連の流れとしてまとめたものである。AF は、USPSTF の概念をもとに¹³⁾、各がん検診の特性を踏まえ作成し、各段階における検討課題を明らかにする。図 2 はいずれのがん検診においても共通となりうる過程と課題を示している。一連の経過は、対象となるがん検診の方法やがんの自然史、治療方法が異なることから、検討対象決定時にモデルを再検討する。

2. 直接的証拠と間接的証拠

死亡率減少効果を示す指標は、直接的証拠と間接的証拠に大別される。

直接的証拠とは、AF1 に相当し、研究のエンドポイントを最終結果である生存・死亡とし、死亡率減少効果の評価した研究である。無作為化比較対照試験、症例対照研究、コホート研究、地域相関研究などにより、がん検診により死亡率減少効果が証明されたものである。

一方、間接的証拠とは、AF2~8 に相当し、エンドポイントを発見がんなどの中間的結果に設定した研究や検査精度に関する研究が含まれる。AF の各段階を構成する検査精度

(感度・特異度)、発見がんの病期分布、生存率などの研究が該当する。これらは、個々の研究だけでは検討対象となるがん検診による死亡率減少効果を証明することが困難なことから、複数の研究の集積により死亡率減少効果が示唆されるものである。

死亡率減少効果を示す証拠として両者を採用するが、あくまでも直接的証拠が優先する。間接的証拠は、単独ではなく、直接的証拠のある検診方法との比較検討が可能な場合のみ、採用することを原則とする。すなわち、精度の優位性が認められる場合でも、比較対照となる検診方法が無作為化比較対照試験により死亡率減少効果が証明されている必要がある。比較対照となる検診方法の死亡率減少効果が、次善の方法である症例対照研究やコホート研究で証明された場合には、間接的証拠を採用する論拠とはならない。

大腸がん検診においては、便潜血検査化学法が無作為化比較対照試験より死亡率減少効果が証明されている。このため、大腸がん検診で各種の検診方法の有効性を検討する場合、便潜血検査化学法と精度を比較することで、間接的証拠を採用することができる。一方、胃がん検診では、胃 X 線検査の死亡率減少効果が症例対照研究により証明されているが、無作為化比較対照試験は行われていない。このため、胃 X 線検査以外の方法による胃がん検診の有効性を証明するには、各検診方法について直接的証拠が必要となる。

V. 文献検索

1. 検索方法

がん検診による死亡率減少効果の判定には直接的、間接的証拠となりうる論文を対象とし、臨床上の診断や治療の効果に関する論文は対象外とした。検討となるがん検診について作成されたAFをもとに、各段階にそった検討課題を設定し、その課題に対する回答を得ることができる文献を抽出する。

AFを構成する要因のうち、AF1は死亡率減少効果を証明する直接的証拠とし、無作為化比較対照試験、症例対照研究、コホート研究、地域相関研究などを抽出する。AF2～8は、検査精度（感度・特異度）、発見がんの病期、生存率、不利益などの文献を抽出する。検査精度の評価の対象となるのは、検診方法として現在検診に用いられていること、また精密検査についても標準的に行われていることを原則とする。感度・特異度は、真陽性をがんや前がん病変などとし、それを明記していることが必要である。さらに、その算出については、同時法・追跡法などの方法を明らかにしたものとした。なお、がん検診受診に関する障壁（バリア）及び経済評価に関する研究は、死亡率減少効果の検討からは除外し、別途検討する。

MEDLINE、医学中央雑誌を中心に1985年1月から検討時に至る関連文献を抽出する。さらに、EMBASE、CINHALなど他の検索エンジンも追加的に利用する。この他、関連する専門学会誌などは、必要に応じてハンドサーチを行う。

文献の採用・除外条件は、以下のとおりである。ただし、除外条件に適応した論文であっても、必要に応じ追加する場合もある。その判断は専門家の意見に基づき、委員会の協議の上、採否を決定する。

- ① 原著のみを対象とし、総説、その他の報告や資料などは除外する。ただし、メタ・アナリシスを含む系統的総括は検討対象とする。
- ② 学会等のガイドライン関連の論文、国家機関・学会などの年報、統計集等は除外する。
- ③ 原則として無症状の一般集団で検討したものに限定し、有症状者（外来受診者など）は除外する。
- ④ 発見率に関する論文は除外する。
- ⑤ 抄録のないものは除外する

2. 対象文献の選択

文献検索により抽出された論文の抄録から、2人1組となり検討する。2人の評価者は、互いに独立して、個別で評価を行う。文献の採否を検討し、採用する場合にはAFを分類する。ただし、AFの複数の段階に用いられる場合もある。さらに両者の採否の評価結果を照合し、最終的に個別研究評価の対象となる文献を選択する。両者の採否が一致した場合には、そのまま採用・不採用を決定する。採否の判定やAFの分類の不一致例は、評価の

対象となるがん検診の専門家が採否の最終的決定を行う。さらに、採用文献を補足する必要がある場合、久道班報告書第3版²⁾、USPSTFの他、関連学会などのガイドラインの引用文献との照合や専門家からの意見を聴取する。それらの情報をもとに、採用文献の追加を行う場合もある。

VI. 個別研究評価

有効性評価を目的とした個別研究評価の対象は、V. 文献検索で最終確定した文献とする。ただし、経済評価及び検診実施の障害関連の論文は除外し、有効性評価が確立した検診方法について、別途個別研究評価を行う。個別研究評価は、研究方法別のチェック・リストを用いて行う。チェック・リストはスコットランドのSIGN¹⁶⁾及びオーストラリア New South Wales Department of Health のMERGE (Method for Evaluating Research and Guideline Evidence)¹⁷⁾の各臨床ガイドラインの研究評価方法を参照し作成したものである。

各研究方法における問題点を検討し、研究の質を吟味する。チェック・リストは、研究方法別に、系統的総括 (Systematic Review)、無作為化比較対照試験、コホート研究、症例対照研究、検査精度、その他の6種類である (添付書類1・2)。チェック・リストを選択する無作為化割付無しの比較対照試験はコホート研究用チェック・リストを用いる。無作為化比較対照試験、コホート研究、症例対照研究において、検査精度にも言及している文献については、主たる研究の評価に付加し、検査精度の評価も同時に行う。

採用となった文献については、要約を作成する。作成された要約に基づき、エビデンス・テーブルとしてまとめられる。このため、研究方法やAFの分類、対象数やその特性、評価指標とその把握方法、結果をまとめ、評価者のコメントを付記する。

作業の手順は以下のとおりである。

- ① ガイドライン作成委員会の委員2人1組で、対象文献を評価する。2人の評価者は、互いに独立し、個別で評価を行う。
- ② 対象文献の研究方法を確定する。
- ③ 対象文献の著者、論文題名、雑誌、号、ページ、発行年を記入する。
- ④ 対象となるがん検診のAnalytic Frameworkの各段階を示す課題番号を記入する。
- ⑤ 各研究方法別にチェック・リストの項目に、5段階評価 (5, 4, 3, 2, 1)、適応外をチェックし、特記すべき点は、備考欄に記入する。
- ⑥ 各研究の適応率を算出する。

$$\text{適応率 (\%)} = ((\text{各項目のスコア合計}) \times 100) / ((\text{総項目} - \text{非該当項目}) \times 5)$$

- ⑦ 適応率と研究全体の評価 (5段階評価) を参考に、採否を確定する。採否判断は適応率60%以上、研究全体の評価は3以上が一応の目安となるが、評価スコアが低い場合でも、委員会の合議により採用とする場合もある。特に、不利益に関する情報は、適応率や研究全体の評価にかかわらず、採用と判断する場合がある。
- ⑧ 採用の決定した研究の要約を記入する。
- ⑨ ガイドライン作成委員会メンバー2人1組で評価した個別研究のうち、採否の一致しなかった文献については、専門家の意見あるいは委員会の合議により、最終的な採否を決定する。

VII. 証拠のまとめ

採用された文献について、チェック・リストから得られる項目や要約をもとに、各検診方法別にエビデンス・テーブルをまとめる。エビデンス・テーブルに基づき、各検診方法別に、検査の概要と共に、死亡率減少効果を示す直接的証拠と間接的証拠を整理する。検査に伴う不利益を記述すると共に、可能な限り、その発生率にも言及する。

その他の要因としては、死亡率減少効果を示す研究から得られた情報に基づき、対象年齢や受診間隔について可能な範囲で検討する。証拠のまとめには、経済評価やがん検診実施の問題となる障壁（バリア）に関する研究は含まない。

1. 証拠のレベル

検診方法別の直接的証拠及び間接的証拠を統合し、その結果に基づき証拠のレベルを判定する。証拠のレベルは、研究方法及び研究の質から、8段階に分類される（表4）。研究の質については、個別研究の評価となる適応率を参照すると共に、研究数、死亡率減少効果の大きさ、複数の研究が同様の結果を示しているかということも考慮し、判定する。

複数の研究で死亡率減少効果が支持されている場合には、最も信頼性の高い研究方法により証拠のレベルを判断する。証拠の質は、++、+、-の3段階である。研究方法の上で最も信頼性の高い無作為化比較対照試験の評価は、証拠の質を吟味し、1++、1+、1-に分類される。無作為化比較対照試験であっても、無作為割付が不適切である場合など、研究の質に問題がある場合には、1-の判定となり、ガイドラインの推奨の根拠としては採用しない。

系統的総括は、無作為化比較対照試験に限らず、症例対照研究やコホート研究に関するものであっても、1と判定されるが、その質を吟味し、3段階に分類する。

間接的証拠となるAFを構成する研究に基づく評価は、検診から死亡率減少に至る一連の過程で、比較検討できる検査方法について無作為化比較対照試験が行われていることが基本条件となる。さらに質の高い症例対照研究やコホート研究が行われている場合は1+に、検査精度などの研究に限定された場合には2+となる。

たとえば、便潜血検査免疫法は便潜血検査の一手法であり、便潜血検査化学法による無作為化比較対照試験が行われている。さらに、便潜血検査免疫法による症例対照研究により、死亡率減少効果が示唆されており、かつ免疫法の感度は化学法より優れている。以上の点から、便潜血検査免疫法そのものの無作為化比較対照試験は行われていないが、証拠のレベルは1+と判定される。

AFを構成する複数の研究があった場合でも、比較可能な検診方法についての無作為化比較対照試験が行われていない場合、すなわち比較可能な検診方法の直接的証拠がない場合は、死亡率減少効果を判定する証拠のレベルは2-となり、間接的証拠は採用されない。

2. 不利益の評価

がん検診の特性を考慮し、各検診方法別の不利益を比較検討する。がん検診の不利益には、偽陰性、偽陽性、過剰診断、偶発症、放射線被曝、感染、受診者の身体的・心理的負担などが該当する。これらの不利益について、検査方法別の比較表を作成する。偽陰性率、偽陽性率、偶発症などは、可能な限り数値を提示する。特に偶発症の発生率は、わが国における報告を利用する。ただし、報告によっては、単純な比較が困難な場合もあることから、比較表に注釈を付記し、またガイドライン本文における各検査方法の不利益に記載をする。表5に大腸がん検診に関する不利益の比較の具体例を示した。不利益の最終的な評価は、比較表に基づき、委員会の討議で評価する。

Ⅷ. 推奨への翻訳

証拠のレベルを基本に、不利益を考慮した上で、最終的な推奨レベルが決定される(表6)。推奨レベルはAからD及びIの5段階で示されている。経済評価や受診率などの検診実施の障壁(バリア)などに関する研究は推奨の判断基準とはしない。

推奨を示すことで、わが国において集団を対象とした対策型検診と個人を対象とした任意型検診として実施の可否を総括する。

対策型検診は集団を対象としたがん検診で、集団の死亡率減少を目的として実施するものを示し、公共的な予防対策として行われる。一方、任意型検診は、個人の死亡リスク減少を目的とし、個人の任意により受診するがん検診を意味する。

推奨AからCまでは、死亡率減少効果を認め、がん検診としての実施が可能である。ただし、対象が異なり、その段階は重要度を示している。推奨A、B、Cは、基本的な条件が整備された条件下では、受診者の価値観に基づく選択が可能である。

推奨A及びBは、死亡率減少効果を認め、かつ不利益も比較的小さいことから、集団を対象とした予防対策としても、個人を対象とした検診であっても実施可能である。この場合でも、少なからず、検診の不利益は存在することから、実施の際は、インフォームド・コンセントが必要である。

推奨Cは、死亡率減少効果は認められるが、無視できない不利益があるため、集団を対象とした予防対策としての実施は望ましくない。しかし、個人を対象とした検診においては、安全性を確保し、不利益についての十分な説明を行った上での実施は可能である。現状の検査に関する不利益は、医療環境の整備、医療従事者の啓発・教育、インフォームド・コンセントなどの充実により、最小化することは可能である。そのため、環境整備が可能な施設であれば、推奨Cとなる検診方法についても実施できる。

推奨Iは、死亡率減少効果の有無を判断するための研究が不十分なことから、保留となったものが該当する。現段階では死亡率減少効果は不明であることから、個人を対象として検診を行う場合には、その点について十分な説明が必要である。また、現在、実施可能な代替案を示し、その利益と不利益の両者の説明をした上で、受診者の価値観に基づいた意思決定が行えるよう支援する必要がある。ただし推奨Iと判定された検診の実施は、有効性評価を目的とした研究を行う場合に限定されることが望ましい。

がん検診の受診の選択では、適切な情報を提供することにより、受診者がどのような検診を受診するかを判断する情報に基づく意思決定(Informed Decision Making ;IDM)が重要であり、多くのガイドラインにおいて推奨に付記されている。特に、がん検診の有効性が不明であったり、受診者の価値観に影響を与える可能性がある場合には、必須と考えられている¹⁸⁾。

本ガイドラインにおける推奨C及びIがこの状況に合致する。このため、推奨C及びIの場合、受診時に、文書と口頭による説明に基づくインフォームド・コンセントを行う

ことが望ましい。受診前にかん検診に関する情報を提供することで、検診受診の意義や必要性を考慮することが可能になる。ただし、インフォームド・コンセントを得る目的は、医療従事者の保全のためだけではないことに留意する。文書及び口頭による十分な説明の上、可能な限り、受診者個人の価値観に基づく、検査方法を選択できるよう、医療従事者が適切な援助を行うことが必要である^{19) 20)}。

推奨Dは、死亡率減少効果がないという証拠があることから、がん検診として、実施すべきではない。

Ⅸ. ガイドライン・ドラフトの作成

1. ガイドライン・ドラフトの作成

エビデンス・テーブルをもとに、各検診方法別の証拠のまとめを整理し、死亡率減少効果を示す証拠のレベル及び不利益を明らかにする。その結果に基づき、個別の検診方法の最終評価となる推奨レベルを決定する。

ガイドラインでは、作成過程を整理すると共に、可能な限り、実行上の問題点となる経済評価や受診率などの検診実施の障壁（バリア）についても言及する。また、受診者の価値観に基づく検診方法の選択が可能かどうかを検討し、可能な場合には、その支援対策について検討する。単一の検診方法しか推奨されない場合には、受診を動機付けるための支援対策を検討する必要がある。

上記の点を考慮し、ガイドラインには以下の項目が含まれる。

- ① 背景
- ② 目的
- ③ 対象
- ④ Analytic Framework
- ⑤ 文献検索方法
- ⑥ エビデンス・テーブル
- ⑦ 証拠のまとめ
検査の概要・直接的証拠・間接的証拠・不利益・その他の要因（対象年齢・受診間隔など）
- ⑧ 証拠のレベル
- ⑨ 不利益の評価
検査方法別の不利益の対比表を作成する。がん検診の不利益には、偽陰性率、偽陽性率、過剰診断、偶発症、放射線被曝、感染、受診者の身体的・心理的負担などが該当する。
- ⑩ 推奨のレベル
- ⑪ その他の問題点
経済評価や受診率などの検診実施の障壁（バリア）有効性評価研究のための課題などが該当する。
- ⑫ 今後の課題と更新予定
ガイドラインの再評価と、更新の具体的な計画（あるいは有効期限）
- ⑬ 用語の解説
略語・専門用語の解説

2. 経済評価及び検診実施の障壁関連の個別研究評価

がん検診の実施に際して経済性は常に考慮すべき課題であるが、経済評価は、有効性が

確立することが前提である。わが国における経済評価研究が標準化されておらず、このため政策決定の判断材料として不十分な情報しか提供できないことから、本ガイドラインでは推奨の判断基準として用いていない。

有効性の確立した検診方法について、経済評価研究の系統的総括を行い、ガイドラインにおいてその成果に言及するか、別途報告を行う。ただし、がん発見費用などの報告は除外し、費用分析、費用最小化分析、費用効果分析、費用効用分析、費用便益分析など標準化された手法に基づく経済評価研究を対象とする。また、これらの評価は、完全な経済評価（費用最小化分析、費用効果分析、費用効用分析、費用便益分析）であること、さらに、社会の立場あるいは支払い者の立場をとることに加え、最終結果（生存・死亡）による分析が望ましい。検討対象とする経済評価研究のうち、完全な経済評価は、可能な限り Washington Panel²¹⁾ に準拠したものについて検討する。

検診実施上の問題点については、医療資源なども問題に加え、受診率対策など、がん検診の普及・推進に関連する情報をあわせて掲載あるいはガイドラインに付随した情報として提供する。

X. 外部評価

1. 外部評価の手順

ガイドライン・ドラフトを、本研究班に所属しない外部の医療従事者（医師、看護師、保健師など）や、保健行政担当者にガイドライン・ドラフトの評価を依頼する。その際、対象となるがん検診や診断・治療の専門家だけではなく、関連領域となる医療情報、経済評価、看護、地域保健等の幅広い分野からの評価者も含むことを原則とする。大腸がん検診ガイドライン・ドラフトの外部評価は、当該がんの専門家 4 名、関連領域 4 名により行われた。

ガイドライン・ドラフトの評価には、外部評価チェックリスト（添付書類 3）を用いる。このチェック・リストに基づき、外部評価をして頂き、評価者の意見を参考に、委員会内での議論を踏まえ、ガイドライン・ドラフトの修正を行う。また、評価者から寄せられたコメントには、委員会での討議を踏まえて、回答する。

外部評価者から、証拠のレベル、不利益、推奨の変更が必要と判断された場合には、その根拠となる研究を提示して頂く。その研究をもとに、委員会内で再度レビューを行い、変更の必要性について、変更を提案した外部評価者にも参加して頂き、委員会で再検討する。

外部評価で修正が行われた後、がん検診にかかわる研究者、実務担当者、一般の方々などを対象とした公開フォーラムを開催する。公開フォーラム開催前に、ホームページ（科学的根拠に基づくがん検診推進のページ <http://canscreen.ncc.go.jp/>）にガイドライン・ドラフト本文及びその要約を掲載する。

2. 外部評価チェックリスト

外部評価チェックリストは、AGREE のガイドラインチェック・リスト 23 項目^{5) 6)}に加え、がん検診のガイドラインチェック・リストとして新たに開発した。AGREE は EC を中心に開発された臨床ガイドラインのチェックリストである。その構成は、目的、作成手順の規則性、受容性、作成委員会の構成や独立性であり、臨床ガイドラインの評価を目的としている。

本ガイドラインの外部評価には適切ではないと考えられる「ガイドラインの利用者が試験的に用いたか」「ガイドラインは出版前に外部評価を受けたか」の 2 項目は除外した。その結果、本ガイドラインの評価に適切と考えられるのは 21 項目であった。

さらに、COGS でガイドラインの基本条件としている「エビデンス・テーブル及び用語の解説が理解しやすいか」という 2 項目を加えた⁸⁾。独自のチェック項目として、「がん検診の推奨として適切であるか」「わが国において実施可能か」と自由記載欄を追加した。以上の、26 項目について、AGREE と同様に 4 段階評価を行い、問題となる項目には、具体的な意見を求めることとした。

3. 公開フォーラムの開催

外部評価で追加・修正が行われた後、がん検診にかかわる研究者、実務担当者、一般の方々などを対象とした公開フォーラムを開催する。一般公開を行う以前に、ガイドラインの評価、普及などの問題点を明らかにし、その周知を目的としたものである。

公開フォーラム開催前に、ホームページ（科学的根拠に基づくがん検診推進のページ <http://canscreen.ncc.go.jp/>）にガイドライン・ドラフト本文及びその要約を掲載し、公開フォーラム参加者を募集する。さらに、関連学会や行政におけるがん検診の担当者、報道関係者などに、公開フォーラムの開催を周知し、参加者を募る。

公開フォーラムでは、委員会メンバーによる、ガイドラインの目的、評価結果の説明を行い、出席者に質問・意見を求める。その際、証拠のレベルや不利益の評価、推奨の変更が必要と判断された場合には、提案者にその根拠となる研究を提示して頂く。その研究をもとに、委員会内で再度レビューを行い、変更の必要性について、変更を提案した方にも参加して頂き、委員会で再検討する。その他の質問・意見に対して、可能な限り公開フォーラムにおいて回答し、その時点で回答不能な場合には、事後にホームページ上で回答を公開する。

公開フォーラムの概要はホームページに掲載する。また、寄せられたご意見は、ガイドラインの普及や今後の評価の改善に反映させるための資料とする。

X I . ガイドラインの公開と再評価

ガイドラインは、がん検診の提供を検討するすべての関係機関において活用されることを期待して作成される。様々な職種の関係者が容易に内容を理解できるように、本報告以外にも、医療従事者を対象にした普及版、一般向けの解説書、検診受診のためのパンフレットなどを作成すると共に、ホームページ（科学的根拠に基づくがん検診推進のページ <http://canscreen.ncc.go.jp/>）を開設した。医療従事者を対象にした普及版は、医学系雑誌に公表する。ガイドライン普及のために、検診実施の際のインフォームド・コンセントの支援ツールとして、パンフレットや説明用スライドを作成し、ホームページ上での公開を行っていく。

ガイドラインの解説についても、関連学会誌や学会・研修会・講演会などを利用し、がん検診に関わる医療従事者への周知に努めていく。

ガイドライン公開後の評価は、ガイドラインの妥当性や更新時の問題点を明らかにするため、継続的な評価が必要となる。調査法は未だ定型化されていないが、NICEをはじめガイドライン公開後の調査が公表されている^{22) 23)}。これらの結果からは、医療サービスの標準化におけるガイドラインの役割と更新に向けての改善点が示されている。本研究班においても、ガイドラインががん検診の実施にどのように寄与しているか、また推奨に基づいた検診が行われているかについては、ガイドラインの理解や利用に関する調査や官庁統計など各種の公開されている情報を用いて、継続的にモニターし、ガイドライン更新のための情報として利用していく。

XII. 今後の課題

わが国において正しくがん検診を推進していくためには、科学的根拠に基づくガイドラインが必要である。また、その評価方法も国際的に標準化された方法に基づくものでなければならない。本ガイドラインは、最新の知見も含めた対象文献の系統的総括に基づき、死亡率減少効果と不利益に関する科学的根拠を明確にし、わが国において集団を対象とした対策型検診及び個人を対象とした任意型検診の実施の可否について推奨に総括することを目標としている。

しかし、その作成過程には、いくつかの問題点を残している。

ガイドラインでは、対象となる範囲を明確にする必要があるが、わが国においては、がん検診と診療の境界が不明確な点がある。また、推奨の判断基準として、死亡率減少効果ばかりではなく、不利益、実行性、経済性、医療資源などの要因を考慮していくことも検討されなくてはならない。

一方、こうした科学的根拠に基づくがん検診が実施されるような政策支援の仕組みが必要である。これらの問題点については、国際的な動向を見据えながら、さらなる議論を積み重ね、本ガイドラインにおける評価方法の改善・標準化を進めていく予定である。

ガイドラインの利用者

ガイドラインは、がん検診に関連するすべての人々への情報提供を目的としている。特に、がん検診の提供を検討するすべての医療機関において活用されることを期待するものである。わが国におけるがん検診は、市区町村を実施主体とする老人保健事業による公的施策として実施されている。

久道班報告書第3版²⁾では、公的施策として実施されるがん検診を念頭に置いて、そのための判断基準が提供された。わが国においては、公的施策として行われるがん検診以外にも、任意性の高い総合健診や人間ドックなどが少なからず実施されている。どのような実施体制であっても死亡率減少効果の確立したがん検診が優先して提供されるべきである。

本ガイドラインは、日本におけるがん検診に係わるできる限り多くの医療機関において使用できることを目標とした。このため、集団を対象とする対策型検診と、個人を対象とする任意型検診の2つに分けて推奨を示した。しかし、個人を対象としたがん検診と診療の境界を明確にすることが、困難な場合もある。この点も含め、今後ともさらなる議論を通じて、推奨基準の改訂を行う予定である。

不利益の評価

本ガイドラインの推奨のレベルを評価する上で、不利益の評価をどのように行うかについても多くの議論が重ねられた。がん検診の利益である死亡率減少効果の評価方法が整備されている一方で、不利益の評価は曖昧な面が残されている。臨床ガイドラインによっては、バランス・シートを用いた評価を行っている場合もある。

本ガイドラインでは、がん検診の特性を考慮し、各検診方法別の不利益を比較検討する。

がん検診の不利益には、偽陰性率、偽陽性率、過剰診断、偶発症、放射線被曝、感染、受診者の身体的・心理的負担などが該当する。不利益の最終的な評価は、比較表に基づき、委員会の討議で評価した。しかし、偶発症の頻度などは、学会報告などを除いては情報収集が困難であり、正確な情報が入手しにくいばかりでなく、医療技術のレベルが一律ではない臨床の現場の感覚とは乖離する可能性がある。過剰診断については、PSAによる前立腺がん検診のように明らかな問題点が指摘されているものもあるが、一方で、その実態を正確に把握することが困難な場合が多い。同様に、偽陽性率、偽陰性率の定義や算出方法の相違が大きく単純な比較も難しい。こうした点から、不利益の評価を一律にルール化することは困難であり、現状の報告に基づき、委員会内の協議の結果を定性的な記述に基づく最終評価とした。

多くの臨床ガイドラインでは、医療サービスの利益と不利益を推奨の判断基準として採用しているが、その評価の論理は必ずしも明確ではない。USPSTFでは、推奨のための評価グリッドを証拠のレベルと利益と不利益の差で検討する方法を用いている¹³⁾。この方法は、理想的には望ましいと考えられるものの、利益と不利益の差をどのように計測するかという明確な回答をもたらすものではない。AGREEにおいても不利益の望ましい評価例として採用されている、SIGN及びニュージーランドの臨床ガイドラインもその記述のみに留まっている。今後、不利益の評価をどのような形で行うかは、本ガイドラインにおいても継続的な検討課題である。

経済評価

がん検診のガイドラインの作成において死亡率減少効果を明らかにすることはいうまでもないが、政策決定に密接に関わる経済性や検診普及の障壁（バリア）などの問題をどのように扱うかが重要となる。しかし、特に経済評価の重要性は認識されているものの、国際的にもどのような推奨基準に取り入れていくかなどその位置づけは明らかではない。

本ガイドラインでは、経済評価を推奨の判断基準には用いず、有効性の評価された検診方法に付随して検討している。USPSTF¹³⁾、CDC Guide¹⁴⁾、SIGN¹⁶⁾でも同様の対応をとっている。

SIGN¹⁶⁾とNICE¹⁵⁾は英国ガイドラインでありながら、経済評価の扱い方が異なっている。

SIGN¹⁶⁾は、臨床ガイドラインに平行して経済評価を行わないが、すでに公表されている質の高い経済評価研究があれば、その成果を根拠に採用している。また、医療資源についても推奨の判断に必要な場合には、その点についても言及している。

一方、NICE¹⁵⁾では、有効性評価と経済評価の両者を根拠として、医療サービス導入の判断基準としている。このため、ガイドラインの作成には医療経済学者が参画している。判断基準となる経済評価は、NHSの視点から行う、QALY(Quality adjusted life-years)を効果指標とする分析である。現時点で標準的に行われている医療サービスを代替案として、新規の評価対象の増分費用効果比が£30,000/QALY以下であれば、導入可能と判断される。同時に、導入される医療サービスには、地域の状況を考慮した費用や資源の分析を求めて

いる。

有効性評価と共に、経済評価を判断基準とするためには、分析方法を標準化すると共に、明確な判断基準が必要となる。経済評価研究が未整備なわが国においては、経済評価指標として検診のがん発見費用が用いられることが多く、経済評価研究が正しく利用されていない。このため、ガイドラインにどのような形で組み入れていくかについて、また、経済評価研究をどのように利用していくかについては、慎重な対応が必要である。

定期的な更新システムの構築

対象となるがん検診のガイドラインの更新は5年以内を基準とし、各種がん検診における新たな検査方法も含めて、再検討する。

久道班報告書第3版²⁾では、「新たな研究の進展にあわせて、定期的に見直し更新することが重要であり、そのための常設的な機関をわが国に設置することが必要である」と指摘されている。本報告書は、厚生労働省がん研究助成金「がん検診の適切な方法とその評価法の確立に関する研究」班がん研究助成金による報告書であるが、本研究班により定式化されるガイドライン作成方法を参考として、将来的には、国立がんセンター内に常設の委員会を設け、ガイドラインの更新作業を担当することを検討する。

文献

1. 平成 10 年度厚生省老人保健事業推進費等補助金 老人保健福祉に関する調査研究事業
がん検診の有効性評価に関する研究班報告書（主任研究者 久道茂）、財団法人公衆衛
生協会、1998.
2. 平成 12 年度厚生労働省老人保健事業推進費等補助金 がん検診の適正化に関する調査
研究事業 新たながん検診手法の有効性評価報告書（主任研究者 久道茂）、財団法人
公衆衛生協会、2001.
3. Gøtsche PC, Olsen O. Is screening for breast cancer with mammography justifiable? *Lancet*,
2000;355:129-34.
4. AGREE Collaboration Group. Guideline development in Europe: An international comparison.
Int J Technol Assess Health Care.2000; 16: 1039-49.
5. Appraisal of Guidelines, Research and Evaluation in Europe (AGREE) collaboration. AGREE
Instrument, 2003.
6. AGREE Collaboration. Development and validation of an international appraisal instrument for
assessing the quality of clinical practice guideline: The AGREE project. *Qual Saf Health Care*,
2003; 12:18-23.
7. Ollenschlager G, Marshall C, Qureshi S, Rosenbrand K, Burgers J, Makela M, Slutsky J.
Improving the quality of health care: using international collaboration to inform guideline
programmes by founding the Guidelines International Network (GIN).*Qual Saf Health Care*,
2004; 13:455-60.
8. Shiffman RN, Scekelle P, Overhage JM, Slutsky J Grimshaw J, Deshpande AM. Standardized
reporting of clinical practice guideline: A proposal from the conference on guideline
standardization. *Am Intern Med*.2003; 139:493-8.
9. Schunemann HJ, Best D, Vist G, Oxman AD. Letters, numbers, symbols and word: How to
communicate grades of evidence and recommendation. *CMAJ*. 2003; 169:677-80.
10. Grade of Recommendation, Assessment, Development, and Evaluation (GRADE) Working
Group. Grading quality of evidence and strength of recommendation. *BMJ*, 2004; 328:1490-4.
11. U.S. Preventive Services Task Force. Guide to clinical preventive services 2nd edition. Baltimore.
Williams & Wilkins, 1996.
12. Miles A, Cockburn J, Smith RA, Wardle J. A prospective from countries using organized
screening programs. *Cancer*. 2004;101(Suppl.5):1201-13.
13. Harris RP, Helfand M, Woolf SH, Lohr KN, Mulrow CD, Teuysh SM, Atkins D. Current
methods of the U.S. Preventive Services Task Force: a review of the process, *Am J Prev Med*,
2001;20(suppl 3):21-35.
14. Briss PA, Zarra A, Psappaioanou M, Fielding J, Hopkins DP, Woolf SH, Hinmman AR, Harris

- JR. Developing an evidence-based Guide to Community Preventive Services-Methods. *Am J Prev Med.*2000; 18(1S):35-43.
15. National Institute for Clinical Excellence. Guideline development methods. 2004.
 16. Scottish Intercollegiate Guidelines Network. SIGN50: a guideline developer's handbook. Edinburgh: SIGN,2001.
 17. Liddle J, Williamson M, Irwig L. Method for evaluating research and guideline evidence (MERGE).Sydney: New South Wales Department of Health, 1996.
 18. Rimer KR, Briss PA, Zeller PK, Chan ECY, Woolf SH. Informed decision making: What is its role in cancer screening? *Cancer.* 2004;101(Suppl.5):1214-8.
 19. Sheridan SL, Harris RP, Woolf SH. Shared decision making about screening and chemoprevention. *Am J Prev Med.*2004;26 (1):56-66.
 20. Briss P, Rimer B, Reilley B, Coates RC, Lee NC, Mullen P, Corso P, Hutchinson AB, Hiatt R, Kerner J, George P, White C, Gandhi N, Saraiya M, Breslow R, Isham G, Teutsch SM, Hinman AR, Lawrence R. Promoting informed decisions about cancer screening in communities and healthcare systems. *Am J Prev Med.*2004;26(1):67-80.
 21. Gold MR, Siegal JE, Russell LB, Weinstein MC. Cost-effectiveness in health and medicine. New York: Oxford University Press, 1996.
 22. Sheldon TA, Cullum N, Dawson D, Lankshear A, Lowson K, Watt I, West P, Wright D, Wright J. What's the evidence that NICE guidance has been implemented? Results from a national evaluation using time series analysis, audit of patient' notes, and interviews.*BMJ.*2005;329:999-1004
 23. Bahtsevani C, Uden G, Willman A. Outcomes of evidence-based clinical practice guideline: A systematic review. *Int J Tech Assess Health Care.*2004;20:427-33.

図表

図 1 : ガイドライン作成過程

図 2 : がん検診の Analytic Framework と検討課題

表 1 : 久道班第 3 版報告書の評価判定

表 2 : 久道班第 3 版報告書の根拠の質

表 3 : 対策型検診と任意型検診の比較

表 4 : 証拠のレベル

表 5 : 大腸がん検診の不利益

表 6 : 推奨のレベル

図1 ガイドライン作成過程

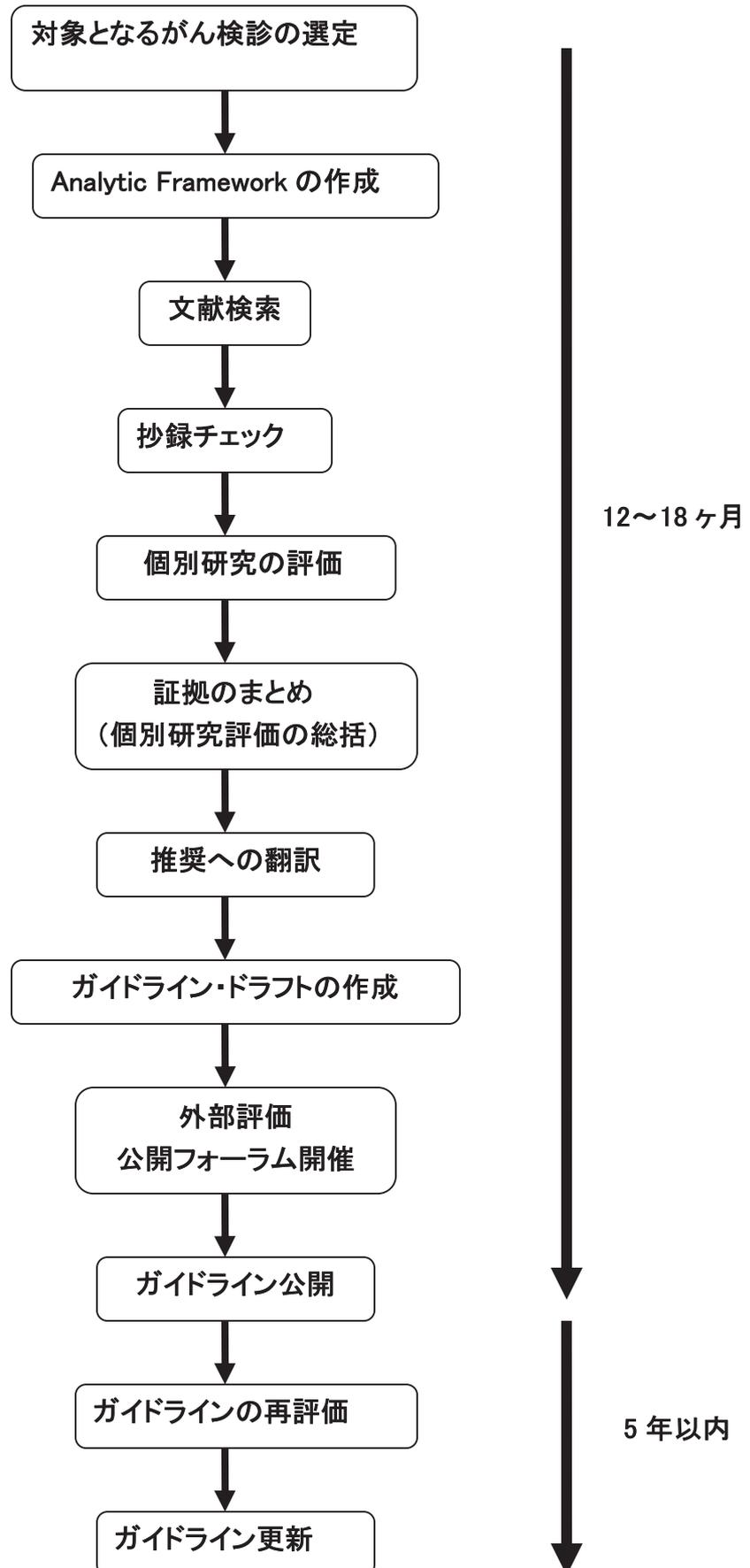
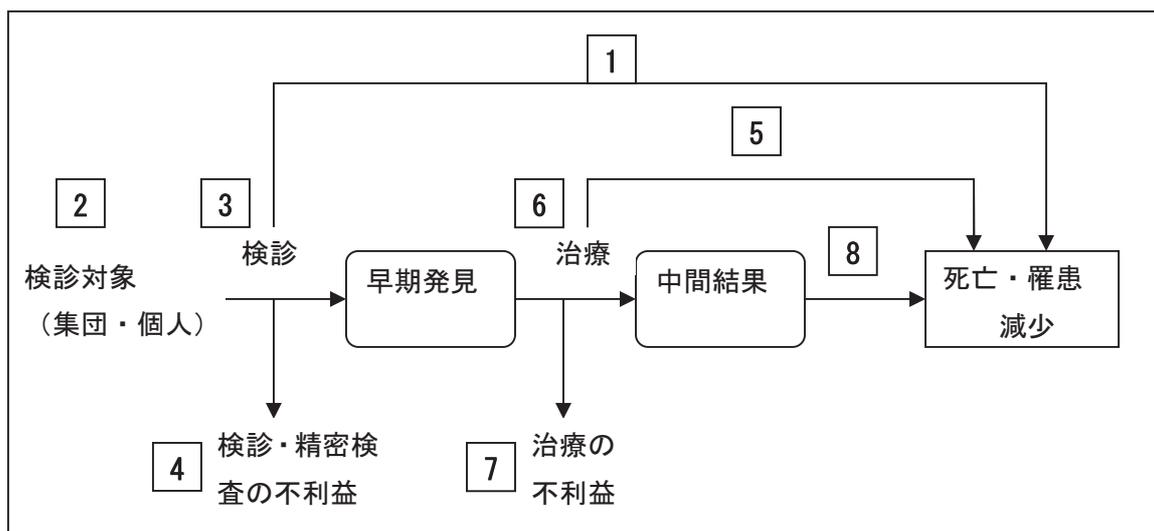


図2 がん検診の Analytic Framework と検討課題



各段階における検討課題

1. がん検診により、対象となるがんの死亡（あるいは罹患）を減少できるか
2. 対象集団における当該がんの罹患率（有病率）
3. 検診
 - 1) 精度（感度・特異度）はどの程度か。どのように算出されているか。他の検診方法と比較可能か。
 - 2) 発見がんの病期分布
4. 検診・精密検査の不利益
 - 1) 偽陰性・偽陽性
 - 2) 偶発症
 - 3) 過剰診断
 - 4) 受診者の負担
5. 適切な治療法が存在し、対象となるがんの死亡（あるいは罹患）を減少できるか
6. 適切な治療法が存在し、中間結果（進行がんなど）を減少できるか
 - 1) 治療効果の評価
 - 2) 検診発見がんと臨床がんとの生存率比較
7. 治療の不利益
 - 1) 偶発症
 - 2) 受診者の負担
8. 中間結果（進行がんなど）の減少が、対象となるがんの死亡（あるいは罹患）の減少につながるか

表1 久道班報告書第3版の評価判定

評価判定	表現
I 群	
I - a	検診による死亡率減少効果があるとすると、十分な証拠がある
I - b	検診による死亡率減少効果があるとすると、相応な証拠がある
I - c	検診による死亡率減少効果がないとすると、相応な証拠がある
I - d	検診による死亡率減少効果がないとすると、十分な証拠がある
II 群	<p>検診による死亡率減少効果を判定する適切な根拠となる研究や報告が、現時点では見られないもの。</p> <p>また、この中には、検査精度や生存率等を指標とする予備的な研究で効果の可能性が示され、死亡率減少効果に関する研究が計画または進められているものを含む</p>

表2 久道班報告書第3版の根拠の質

根拠の質	研究方法
1	無作為割付比較対照試験
2	無作為割付のない比較対照試験
3	コホート研究と症例対照試験
4	地域相関研究と時系列研究
5	その他

表3 対策型検診と任意型検診の比較

検診方法	対策型検診	任意型検診
目的	Organized screening 集団の死亡率を下げる。	Opportunistic screening 個人の死亡リスクを下げる。
検診方法	一定：政府のがん対策担当部局が選ぶ。	様々：個人あるいは個々の検診実施機関が選ぶ。
感度	最も感度の高い検診方法が選ばれないこともある。感度の目標値が設定され、監視される。	通常、最も感度の高い検査が選択される。感度は監視されない。
特異度	高い特異度を選ぶことが、不必要な精密検査による費用や有害事象を避けるために重視される。	特異度が高いことはあまり重要ではない。
受診間隔	一定：適切な費用のもとで、集団の利益が最大化されるように設定される。	様々：がん死亡リスクから個人が最大限守られるように設定される。通常、組織化された検診よりも頻回となる。
検診対象者	一定：特定された年齢範囲の全員	様々：
受診勧奨方法	対象者全員が勧奨される。	一定の方法はなし。
受診機会均等	受診機会が均等になるようにプログラムが計画される。	望ましい。
利益	限られた資源の中で、集団にとって最大化される。	個人にとって最大化される。
不利益	限られた資源の中で、集団にとって最小化される。	最小化は必ずしもされない。

Anne Miles et al, A Perspective from Countries Using Organized Screening Programs. Cancer. 2004;101(5 Suppl):1201-13.

表4 証拠のレベル

証拠レベル	主たる研究方法	内容
1++	無作為化比較対照試験 系統的総括	死亡率減少効果の有無を示す、質の高い無作為化比較対照試験が行われている 死亡率減少効果の有無を示す、質の高いメタ・アナリシス等の系統的総括が行われている
1+	無作為化比較対照試験 系統的総括 AF組み合わせ	死亡率減少効果の有無を示す、中等度の質の無作為化比較対照試験が行われている 死亡率減少効果の有無を示す、中等度の質のメタ・アナリシス等の系統的総括が行われている Analytic Frameworkの重要な段階において無作為化比較対照試験が行われており、 2++以上の症例対照研究・コホート研究が行われ、死亡率減少効果が示唆される
1-	無作為化比較対照試験 系統的総括	死亡率減少効果に関する質の低い無作為化比較対照試験が行われている 死亡率減少効果に関するメタ・アナリシス等の系統的総括が行われているが質が低い
2++	症例対照研究/コホート研究	死亡率減少効果の有無を示す、質が高い症例対照研究・コホート研究が行われている
2+	症例対照研究/コホート研究 AF組み合わせ	死亡率減少効果の有無を示す、中等度の質の症例対照研究・コホート研究が行われている 死亡率減少効果の有無を示す直接的な証拠はないが、 Analytic Frameworkの重要な段階において無作為化比較対照試験が行われており、一連の研究の組み合わせにより死亡率減少効果が示唆される
2-	症例対照研究/コホート研究 AF組み合わせ	死亡率減少効果に関する、質の低い症例対照研究・コホート研究が行われている
3	その他の研究	死亡率減少効果の有無を示す直接的な証拠はないが、Analytic Frameworkを構成する複数の研究がある 横断的な研究、発見率の報告、症例報告など、散発的な報告のみでAnalytic Frameworkを構成する評価が不可能である
4	専門家の意見	専門家の意見

AF: Analytic Framework

注1) 研究の質については、以下のように定義する
 質の高い研究: バイアスや交絡因子の制御が十分配慮されている研究。
 中等度の質の研究: バイアスや交絡因子の制御が相応に配慮されている。
 質の低い研究: バイアスや交絡因子の制御が不十分である研究。

表5 大腸がん検診における不利益

偶発症・受診者の負担	直腸指診	便潜血検査化学法	便潜血検査免疫法	S状結腸鏡検査	全大腸内視鏡検査	注腸X線検査
偽陰性率	報告なし	20.0～75.0%	7.1～70.0%	3.5～4.2% (観察範囲内)	2.5～5.0%	0～20.0%
偽陽性率	報告なし	2.0～20.1%	2.4～30.0%	報告なし	報告なし	報告なし
事前の食事制限	なし	あり(肉類など)	なし	なし～あり	あり(海藻・繊維の多い野菜など)	検査食
事前の薬剤制限	なし	あり(ビタミンCなど)	なし	抗凝固剤	抗凝固剤	なし
薬剤制限による偶発症	-	なし	-	稀だが、出血・血栓症など	稀だが、出血・血栓症など	なし
前処置	なし	なし	なし	洗腸や刺激性下剤	下剤(PEGなど)	塩類下剤(マグコロールなど)
前処置による偶発症	-	-	-	腹痛・吐き気など	腹痛・吐き気など。稀だが、穿孔や腸閉塞。	腹痛・吐き気など。稀だが、穿孔や腸閉塞。
前処置による偶発症(死亡)	-	-	-	報告なし	あり(4人/約11,772万人)	あり(1人/約5,362万人)
前投薬	なし	なし	なし	なし～あり(鎮痙剤・鎮静剤など)	鎮静剤は66%の施設で使用。鎮痙剤も使用される。	高頻度で使用(鎮痙剤)
前投薬による偶発症	-	-	-	ショック・血圧低下・呼吸抑制など	ショック・血圧低下・呼吸抑制など	ショック・血圧低下・呼吸抑制など
前投薬による偶発症(死亡)	-	-	-	可能性あり(前投薬使用の場合)	あり	可能性あり(前投薬使用の場合)
スクリーニング検査偶発症頻度	なし	なし	なし	0～0.0015%(1/65,480)	0.069%(2,038/2,945,518)	0.0051%(4/78,745)
スクリーニング検査偶発症	-	-	-	出血・穿孔など	出血・穿孔など	便秘・穿孔など
スクリーニング検査偶発症(死亡)	なし	なし	なし	なし	0.00088%(26/2,945,518)	報告なし
精密検査の偶発症	あり	あり	あり	あり	あり	あり
感染対策(消毒)	-	-	-	報告はないが、消毒は必要	報告はないが、消毒は必要	-
放射線被曝	-	-	-	-	-	あり(3.5～4.7mSV)

注1) 偽陰性率・偽陽性率の算出方法は、同時法・追跡法などがあるが、その算出条件は研究間で異なる。このため、単純な比較は困難であるが、参考値として、上記表に示している。
(詳細は個別の検査方法の証拠のまとめ参照)

注2) 偶発症の頻度はわが国における報告に基づく
(詳細は個別の検査方法の不利益参照)

表6 推奨のレベル

推奨	表現	対象となる集団(条件)	証拠のレベル
A	死亡率減少効果を示す十分な証拠があるので、実施することを強く勧める。	適切な集団を明確にする (集団及び個人を対象とした検診で行えるなど)	1++/1+
B	死亡率減少効果を示す相応な証拠があるので、実施することを勧める。	適切な集団を明確にする (集団及び個人を対象とした検診で行えるなど)	2++/2+
C	死亡率減少効果を示す証拠があるが、無視できない不利益があるため、集団を対象として実施することは勧められない。 個人を対象として実施する場合には、安全性を確保すると共に、不利益について十分説明する必要がある。	適切な集団を明確にする (個人を対象とした検診で行えるなど)	1++/1+/2++/2+
D	死亡率減少効果がないことを示す証拠があるため、実施すべきではない。	いかなる対象にも行えない	1++/1+/2++/2+
I	死亡率減少効果の有無を判断する証拠が不十分であるため、集団を対象として実施することは勧められない。 個人を対象として実施する場合には、効果が不明であることについて十分説明する必要がある。	実施できる対象を判断する証拠がない	1-/2-/3/4

注1) 集団を対象としたがん検診とは、集団の死亡率減少を目的として実施するものを示し、公共的な予防策として行われる。本ガイドラインは、集団を対象とした対策型検診と定義している。
市町村が行う老人保健事業による集団検診・個別検診や職域の法定健診に付加して行われる検診が該当する。

注2) 個人を対象とした検診とは、Opportunistic Screening、すなわち個人の任意により受診するがん検診を意味する。本ガイドラインは、個人を対象とした任意型検診と定義する。
個人の死亡リスク減少を目的とし、対象となる個人は通常の診療の範囲外となる健康者である。
具体的には、検診センターや医療機関などで行われている総合健診や人間ドックなどに含まれているがん検診が該当する。

注3) 推奨と判定された検診の実施は、有効性評価を目的とした研究を行う場合に限定することが望ましい。

添付書類 1

添付書類 1 チェックリストの解説

I. 概要

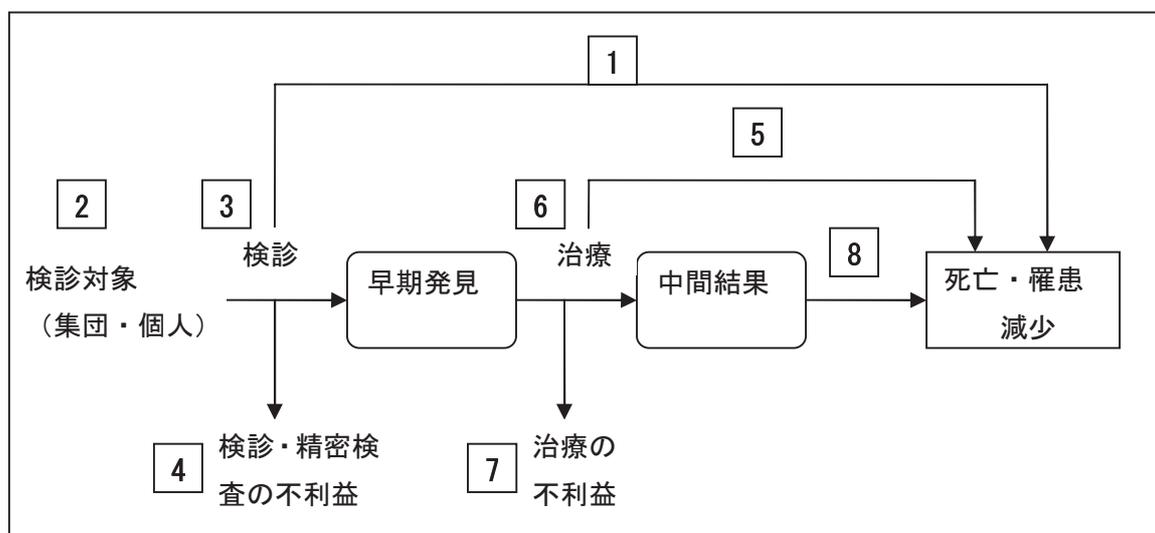
系統的総括(Systematic Review)、無作為化割付比較対照試験(Randomized Controlled Trial、以下 RCT)、コホート研究、症例対照研究、検査精度、その他の 6 研究のチェックリスト・構造化要約がある。その他に分類されるのは、横断研究、前後研究、時系列研究、地域相関研究などである。

6 研究についての共通項目は以下のとおりである。

- 1) 著者、題名、雑誌、号、ページ、発行年
- 2) Analytic Framework

Analytic Framework のどの段階に適応するかを記入する。

図 1. がん検診の Analytic Framework



各段階における検討課題

1. がん検診により、対象となるがんの死亡（あるいは罹患）を減少できるか
2. 対象集団における当該がんの罹患率（有病率）
3. 検診
 - 1) 精度（感度・特異度）はどの程度か。どのように算出されているか。他の検診方法と比較可能か。
 - 2) 発見がんの病期分布
4. 検診・精密検査の不利益
 - 1) 偽陰性・偽陽性
 - 2) 偶発症

- 3) 過剰診断
- 4) 受診者の負担
- 5. 適切な治療法が存在し、対象となるがんの死亡（あるいは罹患）を減少できるか
- 6. 適切な治療法が存在し、中間結果（進行がんなど）を減少できるか
 - 1) 治療効果の評価
 - 2) 検診発見がんと臨床がんとの生存率比較
- 7. 治療の不利益
 - 1) 偶発症
 - 2) 受診者の負担
- 8. 中間結果（進行がんなど）の減少が、対象となるがんの死亡（あるいは罹患）の減少につながるか

3) チェックリストによる評価

5段階（5・4・3・2・1）で評価する。評価は5段階で、5が最良、1が最悪とする。「適応外」については、記載がないものではなく、項目があてはまらにものが該当する。たとえば、RCT 結果の適切性の「対象となる複数施設を比較して、結果の均一性が保たれているか」について、1地域のみの場合、上記には回答不能なので、「適応外」となる。

評価5：十分に記載されている

評価4：記載されている

評価3：ほぼ記載されている

評価2：記載が不十分である

評価1：記載されていない

4) チェックリスト適応率

対象となる研究が、チェック項目をどの程度満たしているかを算出する。

$$\text{適応率 (\%)} = ((\text{各項目のスコア合計}) \times 100) / ((\text{総項目} - \text{非該当項目}) \times 5)$$

5) 採否

適応率と研究全体の評価（5段階評価）を参考に、採否を確定する。採否判断は適応率60%以上、研究全体の評価は3以上が一応の目安となるが、評価スコアが低い場合でも、委員会の合議により採用とする場合もある。特に、不利益に関する情報については、適応率や研究全体の評価にかかわらず、採用と判断する場合がある。

6) 要約

採用の決定した研究の要約をまとめ、記入する。この要約に基づき、各検診方法の評価のためのエビデンス・テーブルが作成される。要約は、最終的にエビデンス・テーブルとしてまとめられる。このため、研究方法やAFの分類、対象数やその特性、評価指標とその把握方法、結果をまとめ、評価者のコメントを付記する。

II. 系統的総括 (Systematic Review)

1. 研究目的

1) 目的が明らかにされているか

問題の焦点を明確にし、そのための適切な評価が行われているかを確認する。

2. 検索方法

1) 文献検索に複数の検索エンジンが用いられているか

文献検索は MEDLINE 単独ではなく、複数の検索エンジン (EMBASE, CINAHL, Cochran Library など) を用いる必要がある。単独の検索による文献抽出の偏りを避けるためである。

MEDLINE は、世界で出版される雑誌の約 25%程度しか含まれておらず、索引できる研究のみを検索するので、限界があることに留意する。また、適宜、対象となるがん検診を対象とする専門誌などのハンドサーチや専門家の意見も加える必要がある。

3. 評価方法

1) 評価方法が明確に記述されているか

系統的総括 (Systematic Review) は、一般的総説と異なり、系統的な評価が行われる。その方法が明確に示されていない場合、証拠の強さは最良「5」ではなく、その質に応じてグレードを下げる。

4. 結果の妥当性

1) 研究対象の選択規準が明確に示されているか

2) 対象となる研究が理論的に統合されているか

3) バイアスをどの程度回避しているか

研究方法や質を考慮した上で、研究対象となる研究の選択規準が明確に示されている必要がある。検討対象の規準が明確であり、対象研究について適切な比較検討が行われているかどうか、を検討する。

対象となる個別研究の質については、各チェックリストを参照する。検討対象となる研究そのものの質が低い場合は、系統的総括によって、過大評価されたり、不適切な結果が導かれる場合がある。個々の研究の包含するバイアスを最小化するために、どのような工夫がなされたかを検討する。3で示された評価方法が適切に統合され、それに基づく結論を導き出しているか吟味する。

5. 結果の適切性

1) 検討結果に基づいて、結論が導きだされているか

2) 検診の利益・不利益の両者を考慮しているか

3) 結果は、わが国の対象に適用できるか

検討結果に従い、結論が導き出されているか、を検討する。検診には死亡率減少効果以外にも、副次的な利益や、無視できない不利益がある。すべての利益と不利益を考慮した結論であるかを、吟味する。さらに、わが国におけるガイドラインの対象にも、同様の結果をあてはめてよいか、どうかを判断する。

Ⅲ. 無作為化比較対照試験 (RCT)

1. 研究目的

1) 目的が明らかにされているか

問題の焦点を明確にし、そのための適切な評価が行われているかを確認する。

2. 対象の設定

1) 無作為割付は適切に行われているか

2) 割付の予測は不可能になっているか

3) 研究開始時、検診群と対照群は類似した集団であるか

4) 割付けられた検診プログラムを除いて、両群の診断・治療の機会は平等であるか

無作為割付が適切に行われ、その方法が明確に記述されていることが基本的条件である。割付が適切ではない場合、評価は「1」となる。また、検診群と対照群（検診受診を奨励しない群）の割付が日付や交互に分配される方法をとっている場合は、適切な無作為化とはいえ、判定は「5」ではなく、その質に応じてグレードを下げる。さらに、割付に際して、検診群・対照群のいずれに割り付けられるかについて、事前の予測が不可能な方法をとるべきである。割付の予測が可能な場合、効果が最大 40%まで過大評価となる報告がある。適切な割付を行うためには、コンピューターシステムの利用などが考えられる。

できる限りバイアスを回避するために、検診群・対照群の両者の属性（性、年齢、人種、社会階層など）は類似した集団であることが前提となる。こうした点は、対象の選択基準を明確にすることで回避することが可能である。両群について、検診受診の有無は無作為に割り付けられるが、その他の診断や治療の機会は平等に与えられている必要がある。追加で受診する検診は、交絡因子となる可能性がある。

3. 対象の管理

1) プロトコールは遵守されているか

2) 追跡期間内の脱落割合は把握されているか

3) 検診群のコンプライアンスは高く保たれているか

4) 対照群のコンタミネーションはないか

5) バイアスをどの程度回避しているか

研究対象者の追跡は、研究当初に計画されたプロトコールを遵守し、特殊の理由がない限り中止や変更はしないことが原則であり、その運用が円滑にすすんでいるかを確認する。

経過観察のための追跡期間内に一定割合の脱落が生じる可能性があるが、その範囲は 20%以内が望ましい。特に、検診群のコンプライアンスは高く保たれている必要がある。追跡期間が長いほど脱落率が高くなることが予想されるが、その理由についても考慮する。研究開始時には、無作為に検診群と対照群に割付けられていても、対照群が検診を受けたり、検診群が検診を受けなかったりということがある。この結果、実際に検診を受けた者のうち、検診群に割付された者の他に、対照群に割付けられた者からも実際に検診を受ける者がでてくる（コンタミネーション）。この結果、検診の効果が過少評価される可能性が

ある。追跡期間内に、こうした可能性はなかったか、を確認する。

これらの可能性を考慮した上で、バイアスがどの程度回避されているかを、総合的に判断する。

4. 結果の把握

1) 結果は、標準的で信頼性や妥当性の保証された方法で把握されているか

結果の把握について、標準的で信頼性や妥当性の保証された方法で把握されているかを検討する。生死の確認については、住民票、死亡小票など、罹患については、がん登録など用いたか、その他の方法の場合、どのような方法によって、結果を把握しているかを吟味する。

5. 解析方法

1) 割付された集団すべてを解析対象としているか (intention-to-treat analysis : 割付重視の原則)

研究当初に割り付けられた検診群・対照群が、追跡期間内で予定通りの過程を経ることはむしろ稀であり、脱落やクロスオーバーが生じる可能性は高い。しかし、解析にあたっては、検診群・対照群に割り付けられた全例を検討対象とする、intention-to-treat analysis (割付重視の原則) を用いているか、検討する。検診群に割付けられた人が、たとえ検診を受けなくても検診群として取り扱う。脱落やクロスオーバーの症例を意図的に除外せず、当初割り付けられた群間の比較を行うことが、解析の基本条件である。

6. 結果の適切性

1) 対象となる複数施設（地域）を比較して、結果の均一性は保たれているか

2) 検診の利益・不利益の両者を考慮しているか

3) 結果は、わが国の対象に適用できるか

複数施設（地域）を対象とした多施設共同研究の場合、施設間の結果に相違がないかどうかを検討する。

検診に関するすべての利益と不利益をも考慮した結論であるかを、吟味する。さらに、わが国におけるガイドラインの対象にも、同様の結果をあてはめてよいか、どうかを判断する。

IV. コホート研究

1. 研究目的

1) 目的が明らかにされているか

問題の焦点を明確にし、そのための適切な評価が行われているかを確認する。

2. 対象の設定・管理

1) 対象群の抽出について明確な規準が設定されているか

2) 研究対象群のうち、各群について解析対象となる集団の割合に相違はないか

3) 研究開始時に、対象者が当該がんの診断を受けている可能性はないか

4) 追跡期間内の脱落割合は示されているか

ある集団から、研究対象を抽出する場合、もとの集団の特性を把握した上で、そこからどのような集団が抽出されたかの規準を明確にする必要がある。また、抽出された集団間の特性を比較検討することが望ましい。研究対象群のうち、各群（検診受診群・非受診群）について解析対象となる集団が、もとの集団からみてどの程度の割合を占めているかを算出できるか、両群が類似の集団であるか、を確認する必要がある。また、研究当初に、検診非受診群にすでにごんと診断された例が紛れ込んでいないかを確認する。すでに「がん」と診断されていた例が紛れ込むと、最終的な結果にも影響を与える。その可能性を留意し、解析に際して、感度分析などの方法がとられているかも、あわせて検討する。

経過観察のための追跡期間内に一定割合の脱落が生じる可能性があるが、その範囲は20%以内とする。追跡期間が長いほど脱落率が高くなることが予想されるが、その理由についても考慮する。

解析対象は、抽出もととなった集団にとって、代表性のある集団であることが望ましい。追跡可能であった対象者と途中脱落者について、検診受診の有無別に比較し、両者の相違が明確になった場合には、結果の評価には慎重を期する。

3. 結果の把握

1) 評価指標となる結果は明確に定義されているか

2) 結果は信頼性の高い方法で把握されているか

評価指標とする結果を明確に規定しているかを確認する。さらに、評価のエンドポイントを設定、追跡期間内における、結果を正確に把握する必要がある。結果の把握について、標準的で信頼性や妥当性の保証された方法で把握されているかを検討する。生死の確認については、住民票、死亡小票など、罹患については、がん登録など用いたか、その他の方法の場合、どのような方法によって、結果を把握しているかを吟味する。

4. 受診歴の把握

1) 検診受診歴が適切に把握されているか

2) 受診歴が、結果に影響を及ぼす可能性はないか

3) 検診受診歴は、複数回の受診歴があるかを確認しているか

検診の評価に影響する検診受診歴を正確に把握し、受診歴の有無により比較検討が可能

であるか否かを確認する。また、検診の受診歴が一回だけではなく、複数回の受診歴があるかを確認することが望ましい。すなわち、研究開始時に初回の受診歴のみの把握では不十分なことから、追跡期間内での検診受診がすべて把握されているかを確認する。初回のみの把握の場合には、判定を1段階低下させる。

5. バイアス・交絡因子への対処

- 1) 信頼性、妥当性のある評価方法が用いられているか
- 2) 主たるバイアス・交絡因子を明らかにし、研究デザインや解析の上で適切に対処しているか

研究の質を高めるためには、信頼性、妥当性のある評価方法が用いられていることが必要である。また、主たるバイアス・交絡因子を明らかにし、研究デザインや解析の上で適切に対処している必要がある。

6. 解析方法

- 1) 信頼区間は示されているか
- 2) 多変量解析の適合度は示されているか

統計的検証を行った場合、その方法を明確にし、信頼期間を示す必要がある。多変量解析が行われた場合には、適合度を示しているか確認する。

7. 結果の適切性

- 1) 検診の利益・不利益の両者を考慮しているか
- 2) 結果は、わが国の対象に適用できるか

検診に関するすべての利益と不利益をも考慮した結論であるかを、吟味する。

行われた研究が、前向きか後ろ向きかを確認し、後ろ向きの場合は、証拠の強さの段階を落とす必要がある。さらに、わが国におけるガイドラインの対象にも、同様の結果をあてはめてよいか、どうかを判断する。

V. 症例対照研究

1. 研究目的

1) 目的は明らかにされているか

問題の焦点を明確にし、そのための適切な評価が行われているかを確認する。

2. 対象の選択

1) 症例群と対照群は、同一の集団から抽出されているか

2) 症例群と対照群の除外規準は同一であるか

3) 症例群は明確に定義され、対照群と識別されているか

4) 対照群は、症例ではないことが確認されているか

5) 症例群と対照群の診断の機会が平等であるか

6) 症例群と対照群は、抽出のもととなる集団に占める割合が示されているか

7) 解析対象者と非対象者について、類似性や相違点を比較しているか

症例群と対照群は、同一の集団から抽出されていることが原則であり、抽出もととなる集団について明確に定義されている必要がある。また、症例群と対照群の除外規準は同一でなくてはならない。除外規準が不明確な場合には、バイアスを回避できない可能性がある。

症例群は明確に定義され、対照群とは識別されている必要がある。症例群は、抽出もととなる集団からみて、代表性があることが望ましい。また、対照群は、症例ではないことが確認されている必要がある。症例群と対照群の受診歴は、同様の手順に沿って収集された情報であることが望ましい。また、診断の機会について平等であることを確認する。

解析対象となる症例群と対照群が抽出もととなる集団に占める割合が示されているかを確認する。解析対象となる割合が低い場合や両群の相違が著しい場合は、研究結果の信頼性を損ねる可能性がある。解析対象は、抽出もととなる集団からみて、代表性があることが望ましい。

3. 受診歴の調整

1) 症例群の確定に受診歴の影響がないような対策がとられているか

2) 受診歴は、標準的で妥当性・信頼性の高い方法で把握されているか

症例群を確定する際、受診歴の影響を受けないようにしているか確認する。症例群と対照群の受診歴は、標準的で妥当性・信頼性の高い方法で把握されていることが望ましい。

4. バイアス・交絡因子の対処

1) 主たるバイアス・交絡因子を明確にし、研究デザインや解析上で適切に対処しているか

研究の質を高めるためには、信頼性、妥当性のある評価方法が用いられていることが必要である。また、主たるバイアス・交絡因子を明らかにし、研究デザインや解析の上で適切に対処している必要がある。

5. 解析方法

- 1) 信頼区間は示されているか
- 2) 多変量解析の適合度は示されているか

統計的検証を行った場合、その方法を明確にし、信頼期間を示す必要がある。多変量解析が行われた場合には、適合度を示しているか確認する。

6. 結果の適切性

- 1) 検診の利益・不利益の両者を考慮しているか
- 2) 結果は、わが国の対象に適用できるか

検診に関するすべての利益と不利益をも考慮した結論であるかを、吟味する。さらに、わが国におけるガイドラインの対象にも、同様の結果をあてはめてよいか、どうかを判断する。

VI. 検査精度

1. 研究目的

1) 目的が明らかにされているか

問題の焦点を明確にし、そのための適切な評価が行われているかを確認する。

2. 標準的方法との比較

1) 信頼性に高い標準的方法と比較しているか

2) 検討対象となる検査と標準的方法は、各々独立して（相互の結果はブラインドで）測定されているか

3) 標準的方法と検討対象となる検査相互の影響がないように実施されているか

新たな検査法の精度を評価するためには、標準的に行われている検査法との比較検討を行う。標準的検査はすでに精度が明らかにされていることから、新たな検査法の精度を評価する基準点となりうる。その点からも、標準的方法自体が、信頼性の高い科学的根拠に基づいていることが望ましい。

新たな検査法の精度を評価するためには、検討対象となる検査と標準的方法は、各々独立して（相互の結果はブラインドで）、同じ対象について測定されている必要がある。各々の測定結果が、影響しあわないように留意されているかを確認する。

3. バイアスの対処

1) 検討対象となる検査は、確かな研究デザインに従い、比較検討されているか

2) バイアスをどの程度回避しているか

研究の質を高めるためには、信頼性、妥当性のある評価方法が用いられていることが必要である。また、主たるバイアスを明らかにし、研究デザインや解析の上で適切に対処している必要がある。

4. 検査の管理

1) 検査の精度は管理されているか

検査の精度管理が適切に行われ、検査の精度を測定する上で支障をきたす可能性はなかったかを確認する。

5. 結果の適切性

1) 検診の利益・不利益の両者を考慮しているか

検診に関するすべての利益と不利益をも考慮した結論であるかを、吟味する。

2) 結果は、わが国の対象に適用できるか

わが国におけるガイドラインの対象にも、同様の結果をあてはめてよいか、どうかを判断する。

Ⅶ. その他

1. 研究目的

1) 目的が明らかにされているか

問題の焦点を明確にし、そのための適切な評価が行われているかを確認する。

2. 対象

1) 対象群の抽出について明確な規準が設定されているか

2) 検討対象となる検査は、確かな研究デザインに従い、比較検討されているか

3) バイアスをどの程度回避しているか

検討対象となる集団が、どのような規準で選択されたかを明確にしているか、確認する。その際、バイアスをできる限り排除することが望ましい。そのためには、信頼性、妥当性のある評価方法が用いられていることが必要である。また、主たるバイアスを明らかにし、研究デザインや解析の上で適切に対処している必要がある。

3. 結果の把握

1) 結果は信頼性の高い方法で把握されているか

結果の把握について、標準的で信頼性や妥当性の保証された方法で把握されているかを検討する。生死の確認については、住民票、死亡小票など、罹患については、がん登録など用いたか、その他の方法の場合、どのような方法によって、結果を把握しているかを吟味する。

4. 結果の適切性

1) 検診の利益・不利益の両者を考慮しているか

検診に関するすべての利益と不利益をも考慮した結論であるかを、吟味する。

2) 結果は、わが国の対象に適用できるか

わが国におけるガイドラインの対象にも、同様の結果をあてはめてよいか、どうかを判断する。

添付書類 2

チェック・リスト：系統的総括

系統的総括	チェック・リスト・要約	Analytic Framework	全体の評価 (5.4.3.2.1)	チェック・リスト・適合率	採否
著者					
題名					
雑誌					
号・ページ					
発行年					
チェック・リスト	小項目	評価 (5.4.3.2.1)	適応外	コメント	
研究目的	目的が明らかにされているか				
検索方法	文献検索に複数の検索エンジンが用いられているか				
評価方法	評価方法が明確に記述されているか				
結果の妥当性	対象とする研究の選択頻度が明確に示されているか 対象とする研究の選択基準が明確に示されているか バイアスなどの程度回避しているか				
結果の適切性	検討結果に基づいて、結論が導きだされているか 検診の利益・不利益の両者を考慮しているか 結果は、わが国の対象に適用できるか				
要約	付加事項	内容			
対象となる研究方法					
検診方法					
対象集団の特性	性、年齢、有病率など				
対象集団の設定条件	地域性、医療機関など				
評価指標					
結果	概要(統計的有意性も含め)				
不利益	偶発症、受診者の負担など				
研究全般に関するコメント					

子エック・リスト: 無作為化比較対照試験

無作為化比較対照試験	子エック・リスト・要約	Analytic Framework	全体の評価 (5.4.3.2.1)	子エック・リスト適応率	採否
著者					
題名					
雑誌					
号・ページ					
発行年					
子エック・リスト	小項目	評価 (5.4.3.2.1)	適応外	コメント	
研究目的	目的が明らかにされているか				
対象の設定	無作為割付は適切に行われているか 割付の予測は不可能になっているか 研究開始時、検査群と対照群は類似した集団であるか 割付けられた検査プログラムを除いて、両群の診断・治療の機会は平等であるか				
研究対象の管理	プロトコルは遵守されているか 追跡期間内の脱落割合は示されているか 検査群のコンプライアンスは高く保たれているか 対照群のコンタミネーションはないか バイアスをどの程度回避しているか				
結果の把握	結果は、標準的で信頼性や妥当性の保証された方法で把握されているか				
解析方法	割付された集団すべてを解析対象としているか (intention-to-treat analysis : 割付重視の原則)				
結果の適切性	対象となる複数施設(地域)を比較して、結果の均一性は保たれているか 検査の利益・不利益の両方を考慮しているか 結果は、わが国の対象に適用できるか				
要約	付加事項				
検査方法	検査群/非検査群				
対象数	性、年齢、有病率など				
対象集団の特性	地域性、医療機関など				
対象集団の設定条件					
検査群における受診率・要精検率					
評価指標					
評価指標の把握方法					
結果	概要(統計的有意性も含め)				
不利益	偶発症、受診者の負担など				
研究全般に関するコメント					

子エック・リスト:コホート研究

コホート研究	子エック・リスト・要約	Analytic Framework	全体の評価 (5.4.3.2.1)	子エック・リスト適応率	採否
著者					
題名					
雑誌					
巻:ページ					
発行年					
子エック・リスト	小項目	評価 (5.4.3.2.1)	適応外	コメント	
研究目的	目的が明らかにされているか				
対象の設定・管理	対象群の抽出について明確な規程が設定されているか 研究対象群のうち、各群について解析対象となる集団の割合に相違はないか 研究開始時に、対象者が当該がんの診断を受けている可能性はないか 追跡期間内の脱落割合は示されているか				
結果の把握	評価指標となる結果は明確に定義されているか 結果は信頼性の高い方法で把握されているか				
受診歴の把握	検診受診歴が適切に把握されているか 受診歴が、結果に影響を及ぼす可能性はないか 検診受診歴は、複数回の受診歴があるかを確認しているか				
バイアス・交絡因子への対処	信頼性、妥当性のある評価方法が用いられているか 主たるバイアス・交絡因子を明らかにし、研究デザインや解析の上で適切に対処しているか				
解析方法	信頼区間は示されているか 多変量解析の適合度は示されているか				
結果の適切性	検診の利益・不利益の両方を考慮しているか 結果は、わが国の対象に適用できるか				
要約	付加事項				
検診方法	内容				
対象数	検診群/対照群				
対象集団の特性	性、年齢など				
対象集団の設定条件	地域性、医療機関など				
評価指標					
評価指標の把握方法					
結果	概要(統計的有意性も含め)				
不利益	偶発症、受診者の負担など				
研究全般に関するコメント					

チェック・リスト：症例対照研究

症例対照研究	チェック・リスト・要約	Analytic Framework	全体の評価 (5.4.3.2.1)	チェック・リスト・適合率	採否
著者					
題名					
雑誌					
号：ページ					
発行年					
チェック・リスト	小項目	評価 (5.4.3.2.1)	適応外	コメント	
研究目的	目的は明らかになっているか				
対象の選択	症例群と対照群は、同一の集団から抽出されているか 症例群と対照群の除外規準は同一であるか 症例群は明確に定義され、対照群と識別されているか 対照群は、症例ではないことが確認されているか 症例群と対照群の診断の機会が平等であるか 症例群と対照群は、抽出のもととなる集団に占める割合が示されているか 解析対象者と非対象者について、類似性や相違点を比較しているか				
受診歴の調整	症例群の確定に受診歴の影響がないような対策がとられているか 受診歴は、標準的で妥当性・信頼性の高い方法で把握されているか				
バイアス・交絡因子への対処	主たるバイアス・交絡因子を明確にし、研究デザインや解析上で適切に対処しているか				
解析方法	信頼区間は示されているか 多変量解析の適合度は示されているか				
結果の適切性	検診の利益・不利益の両方を考慮しているか 結果は、わが国の対象に適用できるか				
要約	付加事項	内容			
検査方法					
対象数	ケース群/コントロール群				
対象集団の特性	性、年齢 有病率 受診率 精神受診率				
対象集団の設定条件	地域性、医療機関など				
評価指標					
評価指標はどのようにして把握されているか					
結果	概要 (統計的有意性も含め)				
不利益					
研究全般に関するコメント					

チェック・リスト:検査精度

検査精度	チェック・リスト:要約	Analytic Framework	全体の評価 (5.4.3.2.1)	チェック・リスト通応率	採否
著者					
題名					
雑誌					
号・ページ					
発行年					
チェック・リスト	小項目	評価 (5.4.3.2.1)	適応外	コメント	
研究目的	目的が明らかにされているか				
標準的方法との比較	信頼性が高い標準的方法と比較しているか 検討対象となる検査と標準的方法は、各々独立して(相互の結果はブラインドで)測定されているか 標準的方法と検討対象となる検査相互の影響がないように実施されているか				
バイアスの対処	検討対象となる検査は、確かな研究デザインに従い、比較検討されているか バイアスをどの程度回避しているか				
検査の管理	検査の精度は管理されているか				
結果の適切性	検診の利益・不利益の両者を考慮しているか 結果は、わが国の対象に適用できるか				
要約	付加事項	内容			
検査方法					
対象数					
対象集団の特性	性、年齢				
対象集団の設定条件	地域性、医療機関など				
評価指標					
評価指標の把握方法					
結果	概要(統計的有意性も含め)				
不利益					
研究全般に関するコメント					

チェック・リスト：その他

その他	チェック・リスト・要約	Analytic Framework	全体の評価 (5.4.3.2.1)	チェックリスト適応率
著者				
題名				
雑誌				
号・ページ				
発行年				
チェック・リスト	小項目	評価 (5.4.3.2.1)	適応外	コメント
研究目的	目的が明らかにされているか			
対象	対象群の抽出について明確な規程が設定されているか 検討対象となる検査は、確かな研究デザインに従い、比較検討されているか バイアスなどの程度回避しているか			
結果の把握	結果は信頼性の高い方法で把握されているか			
結果の適切性	検診の利益・不利益の面を考慮しているか 結果は、わが国の対象に適用できるか			
要約	付加事項	内容		
研究方法				
検査方法				
対象数				
対象集団の特性	性、年齢			
対象集団の設定条件	地域性、医療機関など			
評価指標				
評価指標の把握方法				
結果	概要(統計的有意性も含め)			
不利益				
研究全般に関するコメント				

添付書類 3

外部評価チェック・リスト

平成 13 年に、今後のがん検診を正しく推進するため、がん検診の適正化に関する調査研究事業「新たながん検診手法の有効性評価」報告書(主任研究者 久道茂)が公表されました。

今回、上記の報告書をもとに、最新の知見を含め、新たに胃がん検診ガイドラインを作成しました。本ガイドラインは、各検診方法の死亡率減少効果と不利益に関する科学的根拠を明確にし、わが国において対策型検診・任意型検診としての実施の可否について推奨として総括しています。

本ガイドラインが、がん検診に従事する方々に広く用いられるためには、その内容や記載などに問題がないかを、ガイドライン作成委員会に所属しない方々から、評価していただくことが必要です。ご専門の立場から、ご意見を頂き、公表に先立ち、追加・修正を行った上、ガイドラインを公表する予定です。

このチェック・リストは、臨床ガイドラインのチェック・リストである AGREE Instrument などをもとに作成しております。本チェック・リストにはない項目につきましても、お気づきの点がございましたら、ご意見いただければ幸いです。

外部評価チェック・リストは同封の返信封筒にて、平成 17 年 11 月 5 日(土)までご返送ください。

お手数おかけいたしますが、ご協力のほど、よろしくお願い申し上げます。

平成 17 年 10 月

厚生労働省がん研究助成金による

「がん検診の適切な方法とその評価法の確立に関する研究」研究班

主任研究者 祖父江友孝

問合せ先

国立がんセンター がん予防・検診研究センター 情報研究部 濱島 ちさと

TEL. 03-3542-2511(内)3450 FAX. 03-3547-8587 E-mail: chamashi@gan2.res.ncc.go.jp

1. ガイドラインの目的が明確に記載されているか

【チェック・ポイント】

- A. ガイドラインの目的が、詳細に記述されている。
 - B. ガイドラインにより最良の方法が規定され、リサーチ・クエスチョンに明確に回答している
 - C. 期待される効果が明示されている。
 - 1) 十分に記載されている
 - 2) 記載されている
 - 3) 不十分ながら記載されている
 - 4) 記載されていない
- 3) あるいは4) とお答えの場合、どのように記載すべきかコメントをお願いします。

2. ガイドラインが取り扱う課題が明確に記載されているか

【チェック・ポイント】

- A. 科学的根拠明らかにし、検討課題に回答する手順が明示されている。
 - 1) 十分に記載されている
 - 2) 記載されている
 - 3) 不十分ながら記載されている
 - 4) 記載されていない
- 3) あるいは4) とお答えの場合、どのように記載すべきかコメントをお願いします。

3. ガイドラインを利用する対象が明確に示されているか

【チェック・ポイント】

- A. ガイドラインを利用する対象を明確に定義している
 - 1) 十分に記載されている
 - 2) 記載されている
 - 3) 不十分ながら記載されている
 - 4) 記載されていない

3) あるいは4) とお答えの場合、どのように記載すべきかコメントをお願いします。

4. ガイドラインの対象となる受診者が明確に記載されているか

【チェック・ポイント】

A. 対象となる受診者の条件が明確に記載されている（地域・職域、年齢、性別など）。

- 1) 十分に記載されている
- 2) 記載されている
- 3) 不十分ながら記載されている
- 4) 記載されていない

3) あるいは4) とお答えの場合、どのように記載すべきかコメントをお願いします。

5. 受診者の価値観は考慮されているか

【チェック・ポイント】

A. ガイドライン作成の過程で受診者の立場が考慮されている
B. ガイドライン作成の過程に受診者が何らかの役割を果たしている（作成メンバーに入っている、外部評価を担当しているなど）

- 1) 十分に記載されている
- 2) 記載されている
- 3) 不十分ながら記載されている
- 4) 記載されていない

3) あるいは4) とお答えの場合、どのように記載すべきかコメントをお願いします。

6. 根拠の検索には系統的手法が用いられたか

【チェック・ポイント】

- A. 検索式、検索された情報源と文献検索の日付検索方法の詳細が明示されている。
 - B. 課題に即した検索が行なわれている
 - 1) 十分に記載されている
 - 2) 記載されている
 - 3) 不十分ながら記載されている
 - 4) 記載されていない
- 3) あるいは4) とお答えの場合、どのように記載すべきかコメントをお願いします。

7. 根拠の選択基準が明確に記載されているか

【チェック・ポイント】

- A. 検索された根拠を含めるか除外するか判断する基準が示されている。
 - B. 除外する場合には、その理由が明示されている。
 - 1) 十分に記載されている
 - 2) 記載されている
 - 3) 不十分ながら記載されている
 - 4) 記載されていない
- 3) あるいは4) とお答えの場合、どのように記載すべきかコメントをお願いします。

8. 推奨を導き出した方法は明確に記載されているか

【チェック・ポイント】

- A. 推奨を導き出すのに用いられた方法と、最終結論に達した方法が記載されている。(委員会の合意など)
- B. 評価の不一致があった場合には、どのように対処したかを明示している

- 1) 十分に記載されている
 - 2) 記載されている
 - 3) 不十分ながら記載されている
 - 4) 記載されていない
- 3) あるいは4) とお答えの場合、どのように記載すべきかコメントをお願いします。

9. 推奨を導き出すにあたって、利益、副作用、危険性が考慮されたか

【チェック・ポイント】

- A. 利益、不利益（副作用、危険性）が考慮されている
- B. 利益の大きさについて、他の方法との比較検討を行なっている。

- 1) 十分に記載されている
 - 2) 記載されている
 - 3) 不十分ながら記載されている
 - 4) 記載されていない
- 3) あるいは4) とお答えの場合、どのように記載すべきかコメントをお願いします。

10. 検討された検診方法は現状の胃がん検診の方法を網羅しているか

【チェック・ポイント】

- A. 実施可能な様々な検診方法が考慮されている。
- B. 実施可能なオプションはガイドラインの中に明示されている。

- 1) 十分に記載されている
- 2) 記載されている
- 3) 不十分ながら記載されている
- 4) 記載されていない
- 5) 専門外のため、判断できない

3) あるいは4) とお答えの場合、どのように記載すべきかコメントをお願いします。

11. 推奨とそのもととなった根拠との関連が明確であるか

【チェック・ポイント】

- A. その推奨とそれが基づいている根拠との間に明確な関連がある。
- B. 推奨の根拠となる参考文献のリストが明示されている。
- C. 証拠の有無が明確に記載されている。

A. 胃 X 線検査

- 1) 十分に記載されている
 - 2) 記載されている
 - 3) 不十分ながら記載されている
 - 4) 記載されていない
- 3) あるいは4) とお答えの場合、どのように記載すべきかコメントをお願いします。

B. 胃内視鏡検査

- 1) 十分に記載されている
 - 2) 記載されている
 - 3) 不十分ながら記載されている
 - 4) 記載されていない
- 3) あるいは4) とお答えの場合、どのように記載すべきかコメントをお願いします。

C. ペプシノゲン法

- 1) 十分に記載されている
 - 2) 記載されている
 - 3) 不十分ながら記載されている
 - 4) 記載されていない
- 3) あるいは 4) とお答えの場合、どのように記載すべきかコメントをお願いします。

D. ヘリコバクターピロリ抗体

- 1) 十分に記載されている
 - 2) 記載されている
 - 3) 不十分ながら記載されている
 - 4) 記載されていない
- 3) あるいは 4) とお答えの場合、どのように記載すべきかコメントをお願いします。

12. 胃がん検診の推奨として適切であるか

A. 胃 X 線検査

- 1) 適切である
 - 2) ほぼ適切である
 - 3) 適切ではない
 - 4) 専門外のため、判断できない
- 3) とお答えの場合、どのように修正すべきかコメントをお願いします。また、修正の根拠となる研究を具体的にお示してください。

B. 胃内視鏡検査

- 1) 適切である
- 2) ほぼ適切である
- 3) 適切ではない
- 4) 専門外のため、判断できない

3) とお答えの場合、どのように修正すべきかコメントをお願いします。また、修正の根拠となる研究を具体的にお示してください。

C. ペプシノゲン法

- 1) 適切である
- 2) ほぼ適切である
- 3) 適切ではない
- 4) 専門外のため、判断できない

3) とお答えの場合、どのように修正すべきかコメントをお願いします。また、修正の根拠となる研究を具体的にお示してください。

D. ヘリコバクターピロリ抗体

- 1) 適切である
- 2) ほぼ適切である
- 3) 適切ではない
- 4) 専門外のため、判断できない

3) とお答えの場合、どのように修正すべきかコメントをお願いします。また、修正の根拠となる研究を具体的にお示してください。

13. ガイドラインの記載は理解しやすいか

【チェック・ポイント】

- A. 推奨は具体的であり、あいまいではない。
- B. 推奨は、受診者にとってどのような状況で検診を受けることが適切であるかについて、根拠に基づき具体的に記載されている。

- 1) 十分に理解できる
- 2) 理解できる
- 3) やや理解しにくい
- 4) 理解できない

- 3) あるいは 4) とお答えの場合、どのように記載すべきかコメントをお願いします。

14. 容易に重要な推奨を見分けられるか

【チェック・ポイント】

- A. 利用者が容易に最も妥当な推奨を見つけることができる。
- B. 対象となる検診に関する重要な情報を得ることができる。

- 1) 十分に理解できる
- 2) 理解できる
- 3) やや理解しにくい
- 4) 理解できない

- 3) あるいは 4) とお答えの場合、どのように記載すべきかコメントをお願いします。

15. 専門的な用語についての解説が行われているか

- 1) 十分に記載されている
- 2) 記載されている
- 3) 不十分ながら記載されている
- 4) 記載されていない

3) あるいは4) とお答えの場合、どのように記載すべきかコメントをお願いします。

16. 文献の要約の記載が理解しやすいか

【チェック・ポイント】

- A. 出典などの情報を含んでいる。
- B. 構造化要約あるいはそれに準じた方法で記載されている。

- 1) 十分に理解できる
- 2) 理解できる
- 3) やや理解しにくい
- 4) 理解できない

3) あるいは4) とお答えの場合、どのように記載すべきかコメントをお願いします。

17. 推奨を適用するにあたって考えられる組織の障害が検討されたか

【チェック・ポイント】

- A. 推奨の適用のために必要となる変革などの可能性について議論されている。
- B. ガイドラインを利用・推進するにあたり、障害となる点について検討している。

- 1) 十分に記載されている
- 2) 記載されている
- 3) 不十分ながら記載されている
- 4) 記載されていない

3) あるいは4) とお答えの場合、どのように記載すべきかコメントをお願いします。

18. 推奨の適用に伴う費用に関して考慮されたか

【チェック・ポイント】

- A. 推奨された方法が実施されるために資源の追加が必要となる可能性があるが、資源に対する潜在的影響について検討している。
- 1) 十分に記載されている
 - 2) 記載されている
 - 3) 不十分ながら記載されている
 - 4) 記載されていない
- 3) あるいは4) とお答えの場合、どのように記載すべきかコメントをお願いします

19. 本ガイドラインのための普及の方法が整えられているか

【チェック・ポイント】

- A. 要旨などをまとめた簡略版、教育研修ツール、パンフレット、ホームページなどの情報提供手段が考慮されている
- 1) 十分に整えられている
 - 2) 整えられている
 - 3) やや整えられている
 - 4) 整えられていない
- 3) あるいは4) とお答えの場合、どのように記載すべきかコメントをお願いします。

20. ガイドラインの更新手順は示されているか

【チェック・ポイント】

- A. ガイドライン更新の手順や予定が明示されている。
- 1) 十分に記載されている
 - 2) 記載されている
 - 3) 不十分ながら記載されている
 - 4) 記載されていない

3) あるいは4) とお答えの場合、どのように記載すべきかコメントをお願いします。

21. ガイドラインには監視・監査のための主要な評価基準が示されているか

【チェック・ポイント】

A. ガイドラインが遵守されているかを評価・モニタリングする基準が示されている。

- 1) 十分に記載されている
- 2) 記載されている
- 3) 不十分ながら記載されている
- 4) 記載されていない

3) あるいは4) とお答えの場合、どのように記載すべきかコメントをお願いします。

22. ガイドライン作成委員会のメンバーは、該当する全ての専門家集団からなっているか

【チェック・ポイント】

A. 委員会のメンバー、レビュー担当者、推奨を決定することに関与する人が明示されている。

B. ガイドライン作成グループの構成、専門分野、ガイドライン開発に関連する経験に関する情報が提供されている。

- 1) 十分に記載されている
- 2) 記載されている
- 3) 不十分ながら記載されている
- 4) 記載されていない

3) あるいは4) とお答えの場合、どのように記載すべきかコメントをお願いします。

23. ガイドラインは編集に関してその資金提供源から独立しているか

【チェック・ポイント】

- A. 研究費提供源が明示されている。
- B. 研究費が最終的な推奨に影響しなかったという明確な記載がある。
 - 1) 十分に記載されている
 - 2) 記載されている
 - 3) 不十分ながら記載されている
 - 4) 記載されていない
- 3) あるいは4) とお答えの場合、どのように記載すべきかコメントをお願いします。

24. ガイドライン作成委員会メンバーの利害関係が記録されていたか

【チェック・ポイント】

- A. 全ての委員会メンバーの利害関係が明示されている。
 - 1) 十分に記載されている
 - 2) 記載されている
 - 3) 不十分ながら記載されている
 - 4) 記載されていない
- 3) あるいは4) とお答えの場合、どのように記載すべきかコメントをお願いします。

25. 本ガイドラインは、わが国において実施が可能か

- 1) 十分に実施できる
- 2) 実施可能である
- 3) 実施できる場合もある
- 4) 実施できない
- 3) あるいは4) とお答えの場合、どのように記載すべきかコメントをお願いします。

26. 本ガイドラインについてご意見がありましたら、ご自由にお書きください。

記入日 _____ ご氏名 _____

ご協力ありがとうございました

添付書類 4

用語の解説

ガイドラインにおける用語について解説する。用語の解説は平成 10 年 3 月に公表された厚生省老人保健推進費等補助金老人保健福祉に関する調査研究事業「がん検診の有効性評価班」報告書(主任研究者 久道茂)に基づき、追加・修正を行ったものである。

(以下 五十音順)

アナリティック・フレームワーク analytic framework

対象集団の選定、スクリーニングの実施、偶発症の発生、診断、治療、予後、に至るまでを図式化したもの。個々の事象に関する検討課題(key question)を列挙し、それに沿って研究データを収集し、評価する検診に関する全体の研究データを集約する手法。US Preventive Service Task Force でも、第 3 版から採用され、検診が有効であると評価されるために必要な根拠の論理的な流れや繋がりを図示する。

インテション・トゥ・トリート分析 intention to treat analysis

検診の有効性を評価する無作為化比較対照試験において、割り付けられた検診プログラムをすべての対象者が適切に受診するとは限らない。対照群に割り付けられた対象者が、何らかの理由により、研究群の検診プログラムを受診することやその逆の可能性もある。インテション・トゥ・トリート分析とは、対象者が割り当てられた検診プログラム(あるいは検診を受けないというプログラム)を受診したか否かにかかわらず、割り付けられた群に従って分析することを云う。対象者が方針に従わなかったことも含めて、検診の有効性を評価する第一義的な解析手法である。

エビデンス・テーブル evidence table

収集した文献を、一定の評価基準に沿って、その要約を記載し、データベース化したもの。

オッズ比 odds ratio: OR

症例対照研究において、要因曝露と疾病との関連の強さを評価する指標。検診の効果を評価する場合は「検診受診」を要因曝露とみなし、「当該疾病による死亡」を症例とみなす。オッズ比(OR)が 1.0 以下の場合、要因曝露(検診受診)が疾病(当該がんによる死亡)を予防したことを意味し、1.0 以上では死亡リスクが増加したことを示す。相対危険度の近似値であり、相対危険度と同様の解釈が可能である。すなわちオッズ比(OR)0.3 の場合、検診を受診することで本来1の死亡率が 0.3 に減少した(70%の死亡率減少効果を示した)と表現される。

過剰診断 overdiagnosis

がんによる死亡を防ぐことを目的に、がんによる症状が発現する前に発見し、治療することを目的として、がんの検診が行われる。この前提として、がんは放置すると進行し致命的となると一般的には考えられているが、放置しても、致命的とはならないがんは、一定割合で存在する。すなわち、がんが進行して症状が発現する前に、他の原因で死亡してしまうようながんを早期に発見する場合である。こうした例は、成長速度が極めてゆるやかであったり、

極めて早期にがんを発見した場合、あるいは、がんが発見された人が高齢者であったり重篤な合併症を有する場合に生じやすい。このようながんを診断し、治療することは、受診者にとっての不利益につながることから、過剰診断と呼ばれる。

がん検診の有効性

がん検診の有効性は、がん検診受診者において対象とするがんの死亡率が減少することを第一の指標として評価される。それ以外の指標（発見率、切除率、生存率など）を用いて、がん検診の評価を行う場合にも、わが国の一部では「がん検診の有用性」という表現が用いられることがある。しかし「有用性」という表現は一般的ではなく、諸外国では「有効性」をもって検診の評価が行われる。

感度の測定法(同時法・追跡法)

感度を測定する際に、同一被検者に複数の検査法(感度を求める検査と確定診断に用いる検査)をほぼ同時に行い、確定診断に用いる検査を至適基準(gold standard)として、感度を求める方法を同時法と呼ぶ。一方一つの検査法を行なった後、がん登録などで把握されたその後の罹患を至適基準(gold standard)として、感度を求める方法を追跡法と呼ぶ。検診で発見されない中間期がんを把握できるのは、後者の場合だけである。検査後何年までの罹患を把握するのかわ、値は異なってくる。また、同時法・追跡法とも至適基準(gold standard)と定義するものの精度により、結果の信頼性が影響を受けるため、異なる方法で求められた感度を単純に比較することはできない。

QALY 質調整生存年: Quality Adjusted Life years

経済評価を行う際に、評価するプログラムの結果の指標として用いられる。単純に生存期間の延長を論じるのではなく、生活の質(QOL)を示す効用値で重み付けしたものである。QALY を評価指標とすれば、生存期間(量的利益)と生活の質(質的利益)の両方を同時に評価できる。効用値は完全な健康を1、死亡を0とした上で計測される。たとえばAという治療を受けた場合、5年間生存期間が延長すると仮定し、その間の効用値0.8とすると、QALYは5(年)×0.8=4(QALY)となる。

系統的総括 systematic review

当該分野のオーソリティが文献をまとめて解説をつける総括(Review)では、著者の主観が入り込むことにより、偏った文献の選択や解釈が生じる可能性がある。このため客観的な真理を追究する立場から、一定のルールの下に文献を検索し、批判的吟味を加えてまとめあげられたものを系統的総括と呼ぶ。英国のCochran Libraryは、これを継続的に実施している代表例である。

検診 screening

無症状の者に検査を実施してがんを早期に発見し、早期治療を図ることでその疾患の予後を改善させる(当該がん死亡率を減少させる)こと。スクリーニング検査から精密検査、そしてがんの発見と治療への橋渡しに至る一連の過程およびシステムをいう。

検査の精度 accuracy

検査の目的は、病気のある者とない者とを識別することにある。病気のある者を「陽性」、病気のない者を「陰性」と正しく判定する能力が検査の精度である。具体的には感度、特異度などの指標がある。これらはトレード・オフの関係(トレード・オフの項 参照)にあり、個々に議論するべきではなく、併せて議論しなければならない。

- ・ **感度 sensitivity**

がん検診の場合にはある検査が、がんのある者を「陽性」と正しく判定する割合。下表の中、 $a/(a+b)$ の値である。感度が高いことは、検査法の見落としが少ないことを意味する。

- ・ **特異度 specificity**

ある検査が、がんのない者を「陰性」と正しく判定する場合。下表の中、 $d/(c+d)$ の値である。特異度が高いことは、偽陽性が少ないことを意味し、有病率が低い疾患であるがんを対象とした検診の場合では、最も重要な指標である。

- ・ **偽陰性 false negative: FN**

がんがあるにもかかわらず、検査で「陰性」と判定されるもの。下記の表のbに該当する。見逃し例(interval case)ともいう。偽陰性率は、 $(1-\text{感度})$ として計算される。

- ・ **偽陽性 false positive: FP**

がんがないにもかかわらず、検査で「陽性」と判定されるもの。下記の表のcに該当する。偽陽性率は、 $(1-\text{特異度})$ として計算される。

- ・ **陽性反応適中度 positive predictive value: PPV**

検査で陽性と判定された者における患者の割合である。下表の中、 $a/(a+c)$ の値である。感度と特異性は検査法固有の性能によって決まるのに対して、陽性反応適中度は集団における有病率によっても影響を受けるので、評価指標として用いる場合に留意する必要がある。

		検査	
		陽性	陰性
がん	あり	a	b
	なし	c	d

効果 effectiveness

現実の一般的な集団における有効性の指標。保健医療技術を実際の場に普及させるに当たっては、理想通りにはいかない場合も多くなる。例えば、がん検診の受診を勧められても受診しない者もいれば、検査技術も均一化し難い状況も発生してくる。このような一般的な状況下での有効性を検討するものである。したがって、理想的な条件下での効能は十分であっても効果に乏しいという現象も起こり得る。

効能 efficacy

理想的な条件下で理想的な方法により検診などの保健医療技術が実施された場合の有効

性の指標。例えば、検診の受診を推奨された対象者すべてが検診を受診し、その検査技術も完璧なレベルで行われている状況下での有効性を検討するものである。

効率 efficiency

単位時間または単位費用当たりで得られる効果の指標。得られた効果を、それに要した時間量または費用で割ることによって計算される。

構造化要約 structured abstract

目的・デザイン・方法・結果・結論などの論文の項目と同様の構造で、内容が簡潔にまとめられた抄録。短時間で読むことができ、同様の文献との比較が容易である。

コホート研究 cohort study

ある特性(生活習慣や検診の受診歴など)をもった集団(これを「コホート」という)に対して疾患の罹患や死亡などを追跡することによって、その特性と疾患のリスクとの関連を明らかにする研究。例えば、自発的にがん検診を受診した群と受診しなかった群とで、その後の当該がん死亡率を比較する。無作為化比較対照試験の各群をコホートと表現することもあるが、一般的にがん検診のコホート研究は、受診率の高い集団と低い集団との比較、あるいは、単一コホートの中で受診群と未受診群の死亡率を比較評価するものを指す。

コンタミネーション contamination

汚染、混入などと訳す。がん検診の有効性評価の場合には、無作為化比較対照試験において対照群(検診を受診しないはずの群)に割り付けられた者のうち、実際は検診を受診する者の割合として用いられる。がん検診に関する無作為化比較対照試験の理想は、検診群に割り付けられた者全員が受診し(100%のコンプライアンス)、対照群に割り付けられた者のうちだれも受診しない(0%のコンタミネーション)という状況下での両群の当該がん死亡率の比較である。しかし、各個人の行動を強制することは現実には不可能なので、この2つの問題が起こってくることはやむを得ない。コンタミネーションが高い場合、実際は検診に効果があったとしても、見かけ上の効果(検診群と対照群の死亡率の差)は低く測定される。

コンプライアンス compliance

受容度、応諾率と訳す。対象者のうち、保健行動、医療上の指示を受容する者の割合。例えば、無作為化比較対照試験において検診群に割り付けられた者のうち、実際に検診を受診する者の割合を意味する。コンプライアンスが低い場合、実際は検診に効果があったとしても、見かけ上は十分な効果(検診群と対照群の死亡率の差)が観察されない可能性がある。

時系列研究 time series study

ある集団におけるがん死亡率などの動向について時間の経過をおって観察し、様々な要因との関係を検討する研究手法。例えば、がん検診や新しい治療技術の導入の前後で死亡率が著明に減少していれば、これらの効果を示唆するものといえる。子宮頸がん検診の有効性は主にこの方法により検討されてきた。しかし、他の要因の関与の有無なども慎重に考

慮されなければならない。

至適基準 gold standard

新しい検査法を評価するときに、比較の基準になる方法や手段を指す。たとえば検査法の感度を評価するときに、至適基準として、同時に行った別の検査法(たとえば便潜血検査を評価する場合の全大腸内視鏡検査)とする場合や、がん登録で罹患を把握する場合、あるいは生検組織や剖検診断にする場合がある。至適基準自体の正確性が異なるので、単純に感度だけを比較するのではなく、至適基準をどう定義したのかに着目する必要がある。

死亡率 mortality rate

ある集団内で観察された死亡数を、その集団の観察人年(対象集団一人一人の観察期間の総計)で割ったもの。通常は1年を単位として人口千対、人口10万対等で表す。総死因あるいは各死因別に計算できる。

死亡リスク mortality risk

死亡の危険性を数量的に表すもの。ある要因を持つ者の集団と持たない者の集団(例:検診の受診者と非受診者)との間で、死亡の危険性を比較する場合、死亡に関する相対危険度やオッズ比によって死亡リスクを検討する。

受診率 screening rate

対象集団のうち検診を受診した者の割合。特に、老人保健法による検診の対象者とは、運用上、「それ以外の社会資源や制度をもってしても受ける機会のない住民を対象とするもの」とされている。すなわち、他の医療保険のサービスによる健康診断を受けられる者や医療機関で管理されている者を除外した結果が、ここにおける対象集団に該当する。この定義や計算法は市町村によって異なる。

証拠のレベル level of evidence

個々の検診について、一定のルールに従って、当該検診の有効性に関する証拠の確実性をコード化したもの。本ガイドラインでは1~4に分類し、数字が小さくなるほど証拠の信頼性が高いとした。また1と2については、++、+、-のサブコードをもうけ、同じ1であっても++がもっとも信頼性が高く、-がもっとも信頼性が低いものと定義した。

症例対照研究 case-control study:CCS

過去にさかのぼって後向きに調査する代表的な疫学研究手法で、ある要因の有無による当該疾患罹患あるいは死亡のリスクの違いを検討するもの。がん検診の評価の場合には、当該がんで死亡した症例(あるいは進行がんに罹患した症例)とそうではない対照を一定集団より選択し、両群に対して、がん検診の受診歴や過去の生活習慣、環境曝露に関するデータを収集して、がん検診受診による当該がんによる死亡リスクの低下を検討する。

信頼区間 confidential interval

特定の水準(たとえば95%)の信頼度の中で、真の値を含んでいると考えられる測定値の

範囲。この範囲が狭い場合、計算の結果得られた値の信頼性は高いと考えられる。また、AというデータとBというデータの間で互いの95%信頼区間に重なりがない場合、AとBの間には、統計学的に有意差($p < 0.05$)を示していることと同じ意味である。死亡率減少効果を示す相対危険度やオッズ比の場合、95%信頼区間の上限が1を下回れば、統計学的有意に死亡率減少効果を示したことになる。

推奨 recommendation

ガイドライン作成委員会が、科学的根拠に基づいて判断した、がん検診実施の有無についての勧め。本研究班においては、個々のがん検診について受診すべきか否かという問題についての勧めを「推奨」として示している。

スクリーニング検査 screening test

無症状の者を対象に、疾患の疑いのある者を発見することを目的に行う検査。例えば、胃がん検診の胃部間接X線撮影、大腸がん検診の便潜血検査などが、これにあたる。

生存率 survival rate

ある病気をもつ患者集団において、ある期間までに生存している者の割合。がんでは、一般に5年生存率が治癒率の代わりに用いられることが多い。一定期間まで生死が確認できなかった例(消息不明例や観察打ち切り、観察期間が短いものや、転出など)を打ち切り(censored case)とよび、これらが多い場合は、信頼性に欠ける。がんの臨床研究では、Kaplan-Mayer法を用いた実測生存率(全死因死亡による生存率:observed survival rate)や、補正生存率(当該がん死亡による生存率:cause-specific survival rate)が用いられることが多い。

精密検査 work-up examination

スクリーニング検査で陽性と判定された者を対象に、その疾患の診断を目的に行う検査。例えば、胃や大腸の内視鏡検査や生検(バイオプシー)による組織診などがある。一般にはスクリーニング検査とは異なる検査法であり、スクリーニング検査の単純な再検査(たとえば便潜血陽性の場合の、便潜血の再検査や、喀痰細胞診D以上に対する喀痰の再検査)とは区別しなければならない。

説明と同意、インフォームド・コンセント informed consent:IC

病状や検査・治療方針について、医師等が患者や受診者に対して複数の選択肢があることやその行為による利益と不利益を事前に十分に説明し、患者や受診者自身が理解し納得した上で医療行為を受けてもらうこと。現代の臨床医学では倫理面から不可欠であり、すべての医療行為において、必須とされている。

セルフセレクション・バイアス self-selection bias

検診受診者と非受診者との間で、がんの罹患、死亡及び予後に関するリスク要因の分布が異なることによるバイアスである。発見動機別のがん患者の生存率の比較の際には、一般に、検診受診者は健康意識が高く、喫煙率が低いなどにより、もともと生存率が高いのかも

しれないので、検診発見例の生存率が症状発見例よりも高かったとしても、検診の効果とは言い切れない。また、検診受診者における死亡率減少を検討するための症例対照研究の際には、検診受診者は健康意識が高いために、発病リスク自体が低くなる可能性がある (healthy screenee bias)。逆に、受診者で当該がんの家族歴を有する者が多く含まれるという逆のバイアスが生じる可能性もある。したがって、症例対照研究などの観察的研究において検診受診者と非受診者との間で当該がん死亡率に差があっても、このバイアスのために、それが真に検診効果のみによるものか判別し難いことがある。

相対危険度 relative risk:RR

コホート研究や無作為化比較対照試験において、要因曝露と疾病との関連の強さを評価する指標。ある要因の曝露を受けていない群に対する曝露を受けている群の罹患率(または死亡率)の比として求められる。例えば、あるがん検診を受けた群における当該がんの死亡率が人口 10 万対 30 で、その検診を受けていない群での死亡率が同 50 であった場合、 $30/50=0.6$ が相対危険度 (RR=0.6) となる。この場合、検診受診者では、非受診者よりも当該がんの死亡リスクが 60% になる (40% の死亡率減少効果) と解釈され、検診による死亡率減少効果が定量的に示される。

任意型検診 opportunistic screening

個人を対象としたがん検診とは、個人の任意により受診するがん検診を意味する **Opportunistic Screening** に相当する。本ガイドラインは、個人の受診意志に基づく任意型検診と定義する。任意型検診は、個人の死亡リスクの減少を目的としている。対象となる個人は通常の診療の範囲外となる健常者であり、具体的には、検診センターや医療機関などで行われている総合健診や人間ドックなどに含まれているがん検診が該当する。検診プログラム全体のモニタリングなどは行われていない。

対策型検診 organized screening

事前に定められたガイドラインに従って、組織的に管理して行われる検診。集団を対象としたがん検診として、集団の死亡率減少を目的として実施するものを示し、公共的な予防対策として行われる。対象者や検診間隔、精検方法や治療法が明確に定義されており、受診率、発見率、偽陽性率、偽陰性率などがモニタリングされ、プログラム全体が、適切に運用されているか管理されているものを指す。北欧や英国では、国民のがん死亡減少を目的とした政策として、乳がん検診や子宮頸がん検診の **Organized Screening** が行われている。一方、わが国の老人保健事業によるがん検診は、真の **Organized Screening** とは異なる不完全型がわが国の現状である。本ガイドラインは、集団を対象とした対策型検診と定義する。すなわち、市町村が行う老人保健事業による集団検診・個別検診や職域の法定健診に付加して行われる検診が該当する。

地域相関研究 geographic correlation study

異なる地域集団の間で、ある要因の頻度 (例: 検診の受診率) とある健康現象の頻度 (例: 当該がんの死亡率) との関連を分析することにより、その因果関係を分析評価する研究手法。例えば、あるがん検診の受診率の高い地域ほど当該がんの死亡率が低ければ、その検診

の死亡率減少効果を示唆するものといえる。この結果は理解しやすいものであるが、集団間の交絡因子(性・年齢・喫煙状況)の調整が困難であるため、結果の解釈は慎重に行われなければならない。

致命率 fatality rate

ある病気を持つ患者集団におけるある期間の死亡者の割合。死亡者数をその患者集団の人数で割ることによって計算する。生存率と対をなし、生存率は、(1-致命率)の関係にある。健常者を含めた死亡率と区別する必要がある。

中間期がん interval cancer

一定の間隔でがん検診を実施しているとき、前の検診では陰性と判定されたのにもかかわらず、次の検診が来る前に自覚症状が出現してがんを発見される例のこと。偽陰性例の1つで、一般的には進行速度が速く予後不良である。

トレード・オフ trade off relationship

検査精度の指標は、その検査における正常・異常の基準値によって変わる。検査値が高いほど異常である検査の場合、その基準値を低めに設定すれば、病気のある者のうち陽性と判定される割合は増えるので、感度は上がる。一方、病気がなくても陽性と判定される者(偽陰性)も増えるので、特異度は下がる。逆に、基準値を高めに設定すれば、特異度は上がるが、感度は下がる。このように、一方が上がれば他方が下がるような関係をトレード・オフの関係という。

バイアス 偏り:bias

調査や測定、分析や解釈の過程で、系統的に真の値から離れた結果を生じる誤りのこと。研究の計画実施にあたっては、バイアスの混入しない研究を計画することが重要であり、また研究結果を解釈する際には、バイアスの混入により結果にゆがみが生じていないかの吟味が必要となる。

p 値 p-value

偶然によって差が生じる確率であり、研究データの統計学的信頼性を示す。一般にこの値が5%未満($p < 0.05$ と記載される)の場合、データに「統計学的有意差がある」とし、5%以上(Not significant: NSと記載される)の場合は「統計学的有意差がない」とする。

費用効果分析 cost effectiveness analysis

ある保健・医療サービスの実施に要する費用とそれにより得られる効果(自然単位による健康結果、生存年数の延長など)と比較検討する研究手法。1年生存延長に要する費用(費用効果比)を算出して、同じ効果をもたらす他の保健・医療サービスと比較する。

費用効用分析 cost utility analysis

費用効果分析の1手法であり、効果を効用値の重み付けで調整した単位を用いて、測定する。一般には、生活の質で調整した生存年(QALY)を測定し、費用/効用比を指標とする。

費用便益分析 cost benefit analysis

健康結果を貨幣価値により示したものが便益である。医療サービスに投入された費用と便益を比較検討する方法。例えば、早期発見・治療によって救命できた人が生産活動に復帰することによる経済価値を算出する。

前向き研究 prospective study ・ 後ろ向き研究 retrospective study

研究を立案、開始してから新たに生じる事象について調査する研究を前向き研究、過去の事象について調査する研究を後ろ向き研究と呼ぶ。無作為化比較対照試験は前向き研究の代表的なものであり、症例対照研究は、後ろ向き研究の代表的なものである。検診の有効性評価の場合、前向き研究では、研究開始後に行われた検診を評価し、後ろ向き研究はすでに行われた検診を評価することになる。前向き研究では、交絡因子を事前に把握することで、偏りの制御が可能となるが、研究結果が得られるまでにかかなりの時間を要する。後ろ向き研究では、交絡因子の振り返っての把握が困難なため、偏りの制御が困難であるが、研究は比較的短時間に終了することが可能である。

マッチング matching

複数の集団で予後や要因曝露などを比較する際に、性・年齢などの関連因子を両群間で同一になるように対応させること。例えば、症例対照研究では、各症例に対して、性・年齢などが一致する対照例を選定する。検診発見がんと外来発見がんとで生存率を比較する際には、上記の要因が一致するペアを構成するように症例を収集するなど。これにより関連因子の影響を除外したうえで、検診の効果などをより正確に評価することが可能になる。

無作為化比較対照試験 randomized controlled trial: RCT

予防・治療の効果を科学的に評価するための介入研究。対象者を無作為に介入群(検診など、決められた方法での予防・治療を実施)と対照群(従来通りまたは何もしない)とに割付け、その後の健康現象(罹患率・死亡率)を両群間で比較するもの。ランダム割付比較試験とも呼ばれる。日本語の用語は統一されていないので、Randomized Controlled Trial という英語を略したRCTという用語が使われることが多い。

メタ・アナリシス meta-analysis

一つの研究テーマに対して、複数の研究が行われた場合、結果にバラツキが生じることがある。これらを統合して解析する統計手法である。採用するデータは、系統的総括と同様な一定のルールに基づいて信頼できるものに絞り、それぞれに重み付けをして解析する。信頼性は高いものの、出版バイアス(publication bias:有意な結果のみが発表され、症例数の少ない研究が採用されにくい偏り)などの偏りも生じる。

要精検率 recall rate

スクリーニング検査の結果、精密検査が必要とされたものの割合。基本的にはスクリーニングテストの陽性率(positive rate)と同じである。ただし画像診断や、細胞診などのように、がん以外での精密検査や治療を要すると判断されるものも見つかる場合があり、要精検率を他

の施設や、他の検診と比較する場合、がんを疑ったものだけなのか、非がんを含んだものか基準が異なる場合がある。理想的には要精検率が低く、発見率が高いスクリーニング法が検診としては望ましい。

ラベリング効果 labeling effect

検診の結果を知ることによって受診者が受ける様々な心理的影響。例えば、スクリーニング検査の結果で要精密検査と言われただけで、「自分のがんではないか」と心配し、大変なストレスを受ける者もいる。これを「陰性のラベリング効果」と言う。逆に、検診で正常であったことを知るにより、大いなる安心を得る者もいる。これにより、仕事や日常生活に対する活力が増すという「陽性のラベリング効果」もある。

リードタイム・バイアス lead time bias

検診発見がんと外来発見がんとの間で生存率を比較する際に問題となる偏り。がんの発生から死亡までの時間が両群で等しい(すなわち検診の効果がない)場合でも、早期診断された時間の分(リードタイム)だけ、検診発見がん患者の生存時間は見かけ上長いことになり、したがって見かけ上の生存率もあがることになるという偏りである。生存期間の始点が早期発見の分だけずれるという意味から、ゼロタイム・シフトとも呼ばれる。

利益・不利益 benefit・harm

がん検診の利益としては、集団に対する死亡率減少効果が第一であるが、このほか放置して進行がんで発見された場合に比べて、治療法や医療費が軽減される場合も利益の範疇に入る。一方、不利益としては、見逃し、偽陽性となった人に必要でない精密検査が行われることや、精神的不安を与えること、放射線被曝を被ること、無駄な医療費が必要となることなどである。また、たとえがんであっても、精密検査や治療の結果重篤な偶発症を被ることや過剰診断(overdiagnosis)も不利益の範疇に入る。

罹患率 incidence rate

ある集団においてがん新しく罹患した者の数をその対象集団の観察人年(対象者個人の追跡期間の総和)で割って得られることによって計算する。通常は、人口千人年対あるいは人口10万人年対等(すなわち人口千人あるいは10万人を1年間追跡した場合に新たに罹患する割合)で表す。有病率との(ある時点における、ある集団においてある病気を有する者の割合)違いに注意する。

レンジス・バイアス length bias

がんの成長速度の差によるバイアスである。検診では成長速度の遅いがんのグループが発見される可能性が高くなる(レンジス・バイアス・サンプリングとも呼ばれる)。成長速度の遅いがんは、成長速度の速いがんに比べて一般に予後が良好である。したがって、検診発見がんと外来発見がんとで予後や生存率が異なっていたとしても、それが検診の効果なのか、あるいは各々で発見されるがんの成長速度の違いに由来するものなのか判別し難いというバイアスである。

有効性評価に基づく がん検診ガイドライン作成手順

平成 16 年度 厚生労働省 がん研究助成金

「がん検診の適切な方法とその評価法の確立に関する研究」班

発行日 平成 17 年 3 月 24 日

主任研究者 祖父江 友孝

・表紙デザイン 脇坂 ふじ子
宮下 千代

・印刷・製本 株式会社 三田村印刷所

