

第37回がん検診のあり方に関する検討会	資料2
令和5年1月30日（金）	

令和2年～3年度 厚生労働科学研究費 がん対策推進総合研究事業

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 国立がん研究センター がん対策研究所 検診研究部
中山富雄

背景と目的

- がん検診の有効性評価の指標として、死亡率減少効果が世界標準として用いられてきたが、その成果が得られるのに長い年数を要することから、代替指標がありうるのかを検討する。世界のがん検診ガイドラインで、実際に代替指標での検討があるのか、また代替指標を用いる条件、課題を文献レビューの手法を用いて検討することを目的とする。

海外での予防・検診ガイドラインでの死亡率減少効果以外の代替指標での検討事例

- 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 (USPSTF)、American Cancer Society (ACS)、National Comprehensive Cancer Network (NCCN) 及びInternational Agency of Research for Cancer(IARC)、World Endoscopy Organization(WEO)のガイドラインについて検討した。
- 検討臓器は、子宮頸・大腸・乳房とした。

海外での予防・検診ガイドラインでの死亡率減少効果以外の代替指標での検討実例

臓器	USPSTF	ACS	NCCN
子宮頸がん	浸潤がん罹患率	浸潤がん罹患率	2012から更新されず
大腸がん	感度 test performance (advanced adenoma, 大腸がんの発見率)	感度 罹患率(モデル研究) test performance	感度 test performance
乳がん	死亡率減少効果のみ	死亡率減少効果のみ	死亡率減少効果のみ

子宮頸がん・大腸がんの特性；

- ・ 検診で前がん病変の発見が可能
- ・ 前がん病変からの自然史が明らか
- ・ 前がん病変の切除が容易

子宮頸がん；CIN

→ 円錐切除術

大腸がん；advanced adenoma

→ ポリープ切除術

子宮頸がん検診評価のアウトカム

ランク	アウトカム	
1	子宮頸がん死亡の減少	がん死亡率
2	子宮頸がん罹患（stage IB+の罹患）	浸潤がん罹患率
3	子宮頸がん罹患（微小浸潤がんを含む）	
4	CIN3+病変の減少	
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病変検出増加後のCIN3+病変の減少	
6	スクリーニング陽性の増加 （陽性反応的中度の増加または不変が必要）	

絶対指標

European guideline for quality assurance in cervical cancer screening 2nd edition(2010)

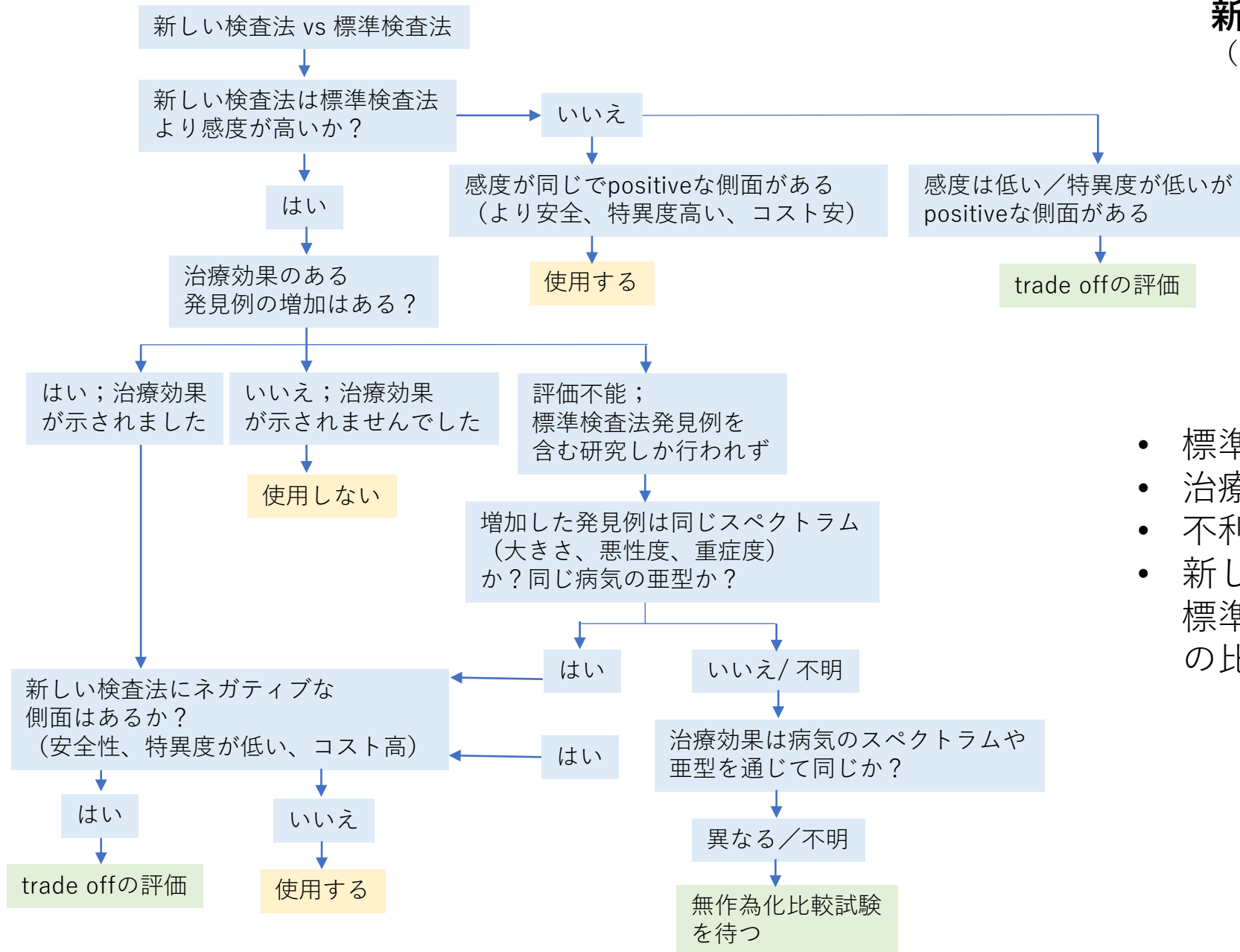
USPSTFの大腸がん検診の評価の変遷

発表年	推奨の内容と評価の変遷	RCT有効性結果発表年	
		便潜血検査化学法	S状結腸鏡
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	1993 US 1996 UK Denmark	
2008	便潜血化学法その他、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（化学法との感度比較がモデルに使われていた）	2008 Sweden	2009 Norway 2010 UK 2011 Italy 2012 US
2016	便潜血単独、便潜血とS状結腸鏡併用法その他、死亡率減少効果が確立していない全大腸内視鏡、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。		

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」と明記されていた。
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ある。

新しい検査法の評価スキーム

(SJ Lord, et al. Ann of Internal Med 2006)



がんの特化していない点
に注意

- 標準検査法と感度の比較
- 治療効果のある発見例の増加
- 不利益の評価
- 新しい検査法での発見症例と標準検査法の発見症例のスペクトラムの比較

WEO/IARC方式

Recommendations for a Step-Wise Comparative Approach to the Evaluation of New Screening Tests for Colorectal Cancer

Graeme P. Young, MD, FRACP, FTSE, AGAF¹; Carlo Senore, MD, MSc²; Jack S. Mandel, PhD, MPH³; James E. Allison, MD, FACP, AGAF⁴; Wendy S. Atkin, MPH, PhD⁵; Robert Benamouzig, MD, PhD⁶; Patrick M. M. Bossuyt, PhD⁷; Mahinda De Silva, MB, BS, FRACP⁸; Lydia Guittet, MD, PhD⁹; Stephen P. Halloran, MBE, FRCPATH¹⁰; Ulrike Haug, PhD¹¹; Geir Hoff, MB, ChB, PhD¹²; Steven H. Itzkowitz, MD, FACP, FACP, AGAF¹³; Marcis Leja, MD, MBA, PhD, AGAF¹⁴; Bernard Levin, MB, BCh, FACP¹⁵; Gerrit A. Meijer, MD, PhD¹⁶; Colm A. O'Morain, MD¹⁷; Susan Parry, MbChB, FRACP¹⁸; Linda Rabeneck, MD, MPH, FRCPC¹⁹; Paul Rozen, MD²⁰; Hiroshi Saito, MD, PhD²¹; Robert E. Schoen, MD, MPH²²; Helen E. Seaman, BSc, PhD²³; Robert J. C. Steele, MD, FRCS²⁴; Joseph J. Y. Sung, MD, PhD²⁵; and Sidney J. Winawer, MD²⁶

BACKGROUND: New screening tests for colorectal cancer continue to emerge, but the evidence needed to justify their adoption in screening programs remains uncertain. **METHODS:** A review of the literature and a consensus approach by experts was undertaken to provide practical guidance on how to compare new screening tests with proven screening tests. **RESULTS:** Findings and recommendations from the review included the following: Adoption of a new screening test requires evidence of effectiveness relative to a proven comparator test. Clinical accuracy supported by programmatic population evaluation in the screening context on an intention-to-screen basis, including acceptability, is essential. Cancer-specific mortality is not essential as an endpoint provided that the mortality benefit of the comparator has been demonstrated and that the biologic basis of detection is similar. Effectiveness of the guaiac-based fecal occult blood test provides the *minimum* standard to be achieved by a new test. A 4-phase evaluation is recommended. An initial retrospective evaluation in cancer cases and controls (Phase 1) is followed by a prospective evaluation of performance across the continuum of neoplastic lesions (Phase 2). Phase 3 follows the demonstration of adequate accuracy in these 2 prescreening phases and addresses programmatic outcomes at 1 screening round on an intention-to-screen basis. Phase 4 involves more comprehensive evaluation of ongoing screening over multiple rounds. Key information is provided from the following parameters: the test positivity rate in a screening population, the true-positive and false-positive rates, and the number needed to colonoscopy to detect a target lesion. **CONCLUSIONS:** New screening tests can be evaluated efficiently by this stepwise comparative approach. *Cancer* 2016;122:826-39. © 2016 The Authors. *Cancer* published by Wiley Periodicals, Inc. on behalf of American Cancer Society. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

KEYWORDS: colonoscopy, colorectal cancer, fecal occult blood test, molecular diagnostics, screening test.

(Cancer 2016)

- 大腸がん検診への新しい検査法を評価するための四段階ステップ
 - これまでは死亡率減少効果がRCTで確認されている方法(FOBT)との精度評価（感度・特異度）に主眼が置かれていた。
 - 感度の測定方法も様々で結果は必ずしも頑健性があると言えないこと、理想的な環境下での一回きりの検査による研究の結果が一般的な検診プログラムに外層できるかは疑問
 - 4段階に分けて、精度評価からリアル・ワールドセッティング（複数回検診）での効果へと、段階を分けて評価。

評価相		第一目的	対象集団と研究手法
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	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集団とがんを持たない集団を対照とする。理想的には診断過程ですべての結果を確認した上での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 特に advanced adenoma を 含む前臨床期病変の発 見精度	発見例が早期癌や 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	単回検診の評価	検診発見病変の性質 偽陽性率、受容性 3.1 検診受診者で精度 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	単回検診で典型的な検診の環境での検査；新しい検査、比較対照（通常の診療の場合も可）がそれぞれ行われる別々のコホートを設定し、受診勧奨からアウトカムまでを追跡。小規模の研究で始めて、発見率を求める大規模研究に移行する。 <u>intention-to-screen</u> 解析が必要。
第4相	複数回検診の評価 （検診プログラムと しての評価）	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	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。 別の検診手法との無作為化比較試験 <u>intention-to-screen</u> 解析が必要

Test performance 研究

- 評価指標を発見率、中間期がん においている。
- 第3相は単回のRCT（狭義のtest performance研究）
- 第4相は複数回検診のRCT(検診プログラムとしての評価)

死亡率減少効果をエンドポイントとした研究の前段階、もっと少ないサンプルサイズ（数千人規模）で行われている。メインは発見率なので、現行の研究は追跡期間が1～2年程度が多い。

（中間期がんについては過小評価？）

第4相は、導入の際に知りたいさまざまな情報

（費用、繰り返し検診参加率、診断までのコンプライアンス、適切な検診間隔、なども評価対象）

第3相 単回検査のRCT

S状結腸鏡を参照値としたネットワークメタアナリシスによるテストパフォーマンス

	アウトカム	S状結腸鏡	S状結腸鏡＋ 便潜血化学法	CTC	全大腸内視鏡
Intention-to-screen解析	AN検出率	1	0.50(0.24-1.00)	0.91(0.49-1.63)	0.96(0.58-1.59)
	CRC検出率	1	0.82(0.23-2.65)	1.51(0.64-3.56)	1.03(0.52-2.05)
Per protocol解析	AN検出率	1	0.70(0.38-1.23)	1.03(0.64-1.58)	1.45(1.00-2.07)
	CRC検出率	1	1.15(0.33-3.66)	1.63(0.71-3.68)	1.49(0.80-2.80)
	参加率	1	0.72(0.40-1.15)	0.91(0.57-1.34)	0.63(0.39-0.93)

25研究、379,556人を招待（各研究あたりの招待数は202～70,096人）

AN; Advanced neoplasia、CRC; Colorectal cancer

Intention-to-screen解析；割り付けられたもの（受診しなかったもの含む）で解析

Per protocol解析；受診したもののみのみで解析

赤字は統計学的有意性を示す

第3相 単回検査のRCT

便潜血免疫法を参照値としたネットワークメタアナリシスによるテストパフォーマンス

	アウトカム	便潜血 免疫法	S状結腸鏡	S状結腸鏡＋ 便潜血化学法	CTC	全大腸内視鏡
Intention-to- screen解析	AN検出率	1	2.29(1.53-3.53)	1.14(0.55-2.30)	2.10(1.18-3.62)	2.20(1.48-3.32)
	CRC検出率	1	1.24(0.78-2.24)	1.02(0.31-3.37)	1.88(0.83-4.61)	1.29(0.77-2.35)
Per protocol 解析	AN検出率	1	2.91(2.21-4.06)	2.04(1.14-3.56)	3.00(0.64-1.58)	4.22(3.16-5.77)
	CRC検出率	1	1.50(1.01-2.69)	1.79(0.54-5.60)	2.48(1.14-5.96)	2.28(1.39-4.07)
	参加率	1	0.73(0.50-0.95)	0.52(0.28-0.84)	0.66(0.39-0.94)	0.45(0.28-0.70)

25研究、379,556人を招待（各研究あたりの招待数は202～70,096人）

AN; Advanced neoplasia、CRC; Colorectal cancer

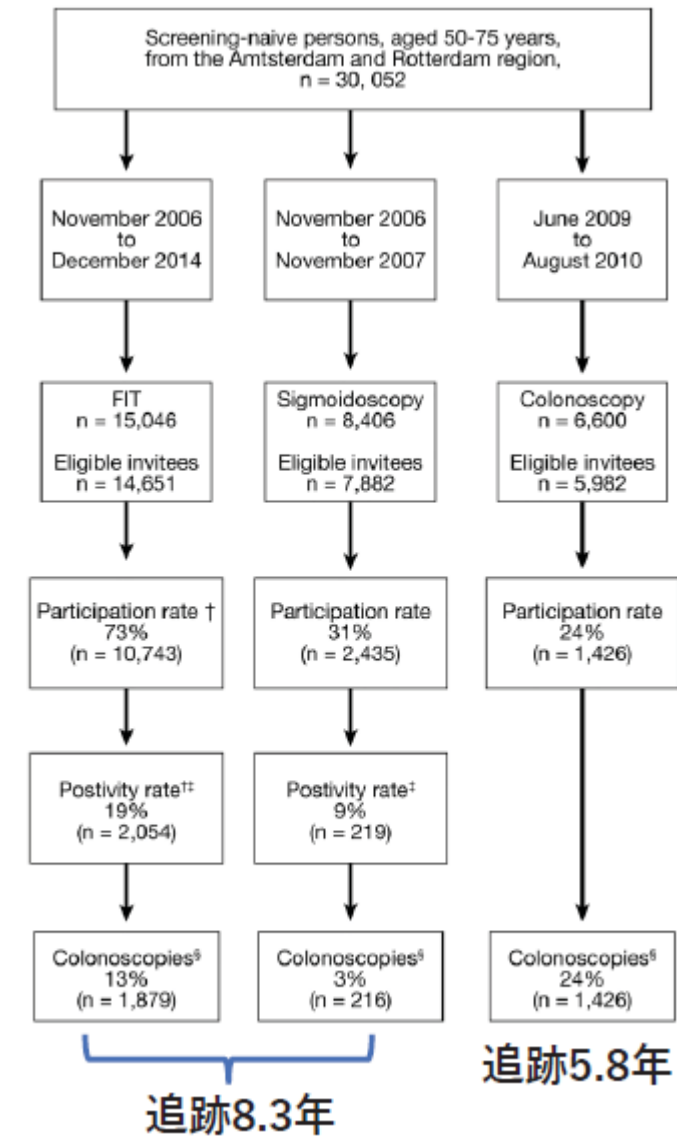
Intention-to-screen解析；割り付けられたもの（受診しなかったもの含む）で解析

Per protocol解析；受診したもののみで解析

赤字は統計学的有意性を示す

第 4 相 複数回検査のRCT

50－75歳 30,052人を 3 群に割付し、便潜血免疫法群は隔年の繰り返し検診、S状結腸鏡・全大腸内視鏡群はベースライン 1 回のみの検診とし追跡した。



	アウトカム	便潜血免疫法 (隔年最大 4 回) %(95%信頼区間)	S状結腸鏡 (1 回のみ) %(95%信頼区間)	全大腸内視鏡 (1 回のみ) %(95%信頼区間)
Intention-to-screen解析	AN検出率	4.5(4.2-4.9)	2.3(2.0-2.7)	2.2(1.8-2.6)
	CRC検出率	0.6(0.5-0.7)	0.2(0.1-0.3)	0.2(0.1-0.3)
Per protocol解析	AN検出率	6.1(5.7-6.6)	7.4(6.5-8.5)	9.1(7.7-10.7)
	CRC検出率	0.8(0.6-0.9)	0.5(0.3-0.9)	0.6(0.3-1.2)
	中間期がん	0.2(0.1-0.3)	0.2(0.1-0.3)	0.01(0-0.1)
	参加率	73%	31%	24%

赤字は統計学的有意性を示す
Grobbee, et al. Clin Gastroenterology, 2020

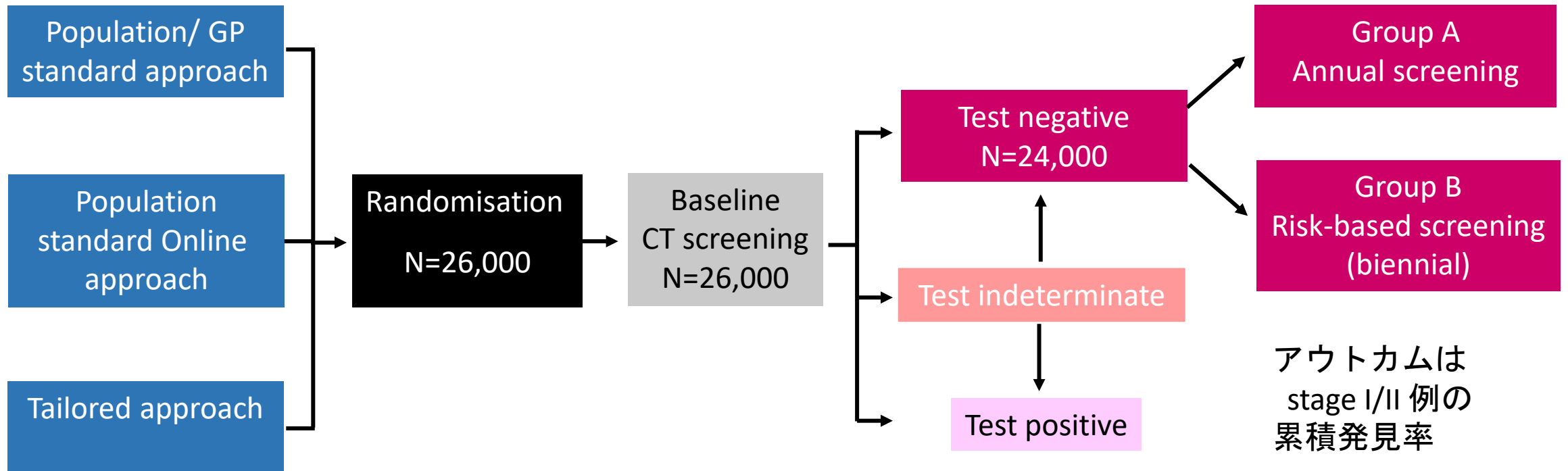
4-IN THE LUNG RUN trial

RECRUITEMENT

RANDOMIZATION

SCREENING

低線量CT検診の検診間隔が1年 vs 2年を評価する無作為化比較試験



アウトカムは
stage I/II 例の
累積発見率

UK, GER, ESP, France, Italy, Netherland

WEO/IARC研究の第4相に匹敵。
Annual screeningは死亡率減少効果がRCTで
証明されている標準法としている。

代替指標の利用条件と混入するバイアス

	浸潤がん罹患率減少	精度評価の積み上げ方式 感度・特異度、test performance、中間期がん
条件	前がん病変が評価する検査法で発見可能であり、その自然歴（放置した場合の浸潤がんの罹患リスク、前がん病変の切除を行った場合の浸潤がんの罹患リスク）が明らかであること。	すでに死亡率減少効果が確認された検診手法が存在すること。その方法との精度比較となる。 これまで評価研究が行われたことがない臓器に対して新たに開発された検診手法に関しては、比較対照となるものがないので、本指標は適応できない。
研究手法	無作為化比較試験 （対照群は無検診あるいは既存の検診手法）	第一段階 後ろ向き精度評価（two-gate;症例対照研究） 第二段階 前向き精度評価（single-gate;コホート研究） 前がん病変発見率／早期がん発見率の比較） 第三段階 既存の検診手法との無作為化比較試験（test performance） 第四段階 既存の検診手法との無作為化比較試験（複数回test performance）
混入するバイアス	過剰診断などのバイアスは混入しにくい。前がん病変を放置した場合浸潤がんに移行する期間を考慮した長い追跡期間を設けないと、効果を過小評価する。	比較する既存の検診手法と新しい検診手法との間で発見がんの性状が大幅に異なる場合、特に前がん病変の発見率が新しい検診手法で飛躍的に高い場合は、過剰診断の影響を受け、過大評価となりやすい。

代替指標を用いる点での注意点

- すでに死亡率減少効果が無作為化比較試験で示されている既存の検査手法との比較という形で試用が始まっており、新しい検査法の評価期間の短縮につながる可能性がある。
- 既存の検査手法自体がないという段階では利用できない。
- 第1相、第2相で新しい検査法での発見がん・前がん病変の特性を既存の検査法と比較することが求められており、過剰診断の可能性について議論が可能。（新しい検査法での発見がんが進行速度の遅いものが多い場合は、死亡率減少効果の評価が必要となる）
- 検診回数や対象年齢などの再評価を行うには適している。