

検査研究の評価・統合に関する研究

研究分担者 寺澤 晃彦 藤田大学 医学部 救急総合内科学 教授

研究要旨

代替え指標として感度・特異度、前がん病変・がん病変を検出する検診パフォーマンス比較が、海外の一部のガイドラインで検討されている。本邦での子宮頸がん検診ガイドライン、大腸がん検診ガイドラインのエビデンスレポート作成段階で行っているこれらの検討を実例として示した。子宮頸がん検診では、既存の方法として医療従事者採取（細胞診、HPV検査）と新しい方法として自己採取HPV検査を比較し、検査結果の一致性、感度・特異度はほぼ同等であることを示した。検診パフォーマンスについてはHPV検査ではどちらが優れるという結果は得られなかったが、細胞診との比較で自己採取HPV検査が有意に高い検診パフォーマンスを示していた。大腸がん検診では、既存の方法として便潜血化学法、新しい方法として免疫法を比較し、感度は免疫法が優れていて、特異度は同等、検診パフォーマンスについては、前がん病変の検出力が免疫法で2.4倍高いものの、がんではほぼ同等であることを示した。わが国のがん検診ガイドライン作成のためのエビデンスの評価においても代替え指標としてのこれらの検討は可能である。ただしその解釈や条件については別途検討の必要がある。

A. 研究目的

がん検診の評価指標として、短期的に得られる指標である感度・特異度や、前がん病変・がん病変を検出する検診パフォーマンス比較が子宮頸がんや、大腸がん検診の評価として海外の一部のガイドラインで行われている。ここでは、「有効性評価に基づく子宮頸がん検診エビデンスレポート」と「有効性評価に基づく大腸がん検診エビデンスレポート」の作成段階で行った新しい検査法の精度評価の実例として示す。

B. 研究方法

1. 子宮頸がん検診

有効性評価に基づく子宮頸がん検診ガイドライン最新版では標準的な検診方法は細胞診単独法、高リスクヒトパピローマウイルス（hrHPV）検診単独法が推奨されている。これらの方法は医療従事者が外来等で検体を採取して検査をする従来型の方法（医療従事者採取）であった。しかし、近年新たな検診方法として検診受診者が自宅等で自ら検体を採取しhrHPV検査を行う方法（自己採取HPV検査）が検討されてきている。医療従事者採取細胞診、医療従事者採取HPV検査を標準的な方

法とし、新たな技術として自己採取HPV検査の精度比較をメタアナリシスの手法を用いて行った。

2. 大腸がん検診

本邦の大腸がん検診は死亡率減少効果が複数の無作為化比較試験で示された便潜血化学法による検診から現在は便潜血ヒト免疫法に移行しており、これまで両者について大腸がん死亡減少効果等の重要な長期健康アウトカムを直接比較したランダム化エビデンスはない。便潜血検査免疫法と化学法の精度比較をメタアナリシスの手法で行った。

（倫理面への配慮）

本研究は既出の文献を用いたシステマティックレビュー、メタアナリシスによる検討であり、倫理面の問題は発生しない。

C. 研究結果

1. 子宮頸がん

1-1. 検査結果の一致率

ペアデザインで医療従事者採取検体と自己採取検体を用い、hrHPV検査による陽性・陰性の判定結果を評価した17件（全14,777人）をhrHPV検査アッセイに依らず検討した。検査の一致率κは

0.45–0.89であり、量的統合解析では統合一致率 κ は0.670 (95%CrI: 0.605–0.737) と比較的高い結果の一致を確認した (図1.1)。陽性、陰性個別の評価では統合陽性一致率 0.840 (95%CrI: 0.760–0.899)、統合陰性一致率0.901 (95%CrI: 0.845–0.939)と採取者が異なっても比較的高い一致度であった。

1-2. 感度・特異度

次に、コルポスコピー下生検により、前がん病変である子宮頸部上皮内腫瘍異形度2以上 (CIN2+) を評価目的病変として検診精度を算定したペアデザイン研究12件を評価した。コルポスコピーおよび生検の適応は各研究のアルゴリズムにより、確認バイアスについては補正していない。医療従事者採取検査の感度・特異度を基準 (1.0) として、自己採取HPV検査の相対的な感度・特異度 (相対感度・特異度) を求めた結果では、相対感度は点推定値で0.75~1.20、量的統合解析での統合相対感度は0.842 (95%CI: 0.702–0.958) であり、自己採取HPV検査の感度はやや低くなっていた。一方、相対特異度の点推定値は0.96~1.02であり、統合相対特異度は0.982 (95%CI:0.966–0.998) とほぼ同等の結果であった (図1.2)。

1-3. CIN2+を検出する1ラウンド検診パフォーマンスの比較)

ランダム化比較で医療従事者採取検体による標準的な検診法と自己採取検体でのhrHPV検査でCIN2+を検出する1ラウンド検診パフォーマンスを比較した研究 (無作為化比較試験) を評価した。医療従事者採取検体による細胞診検査と自己採取検体でのhrHPV検査を比較した無作為化比較試験が11件 (96,161人)、医療従事者採取検体hrHPV検査と自己採取検体hrHPV検査を比較した無作為化比較試験3件 (39,188人) を対象とした。無作為化比較試験の研究対象者集団はルチン検診対象者 (高い参加率が期待される集団) から従来の検診不参加 (参加率は低い集団) と多様であり、中央値34%、範囲9%~94%であった。この報告ではランダム化による集団間の比較同等性が必ずしも担保されないことを許容しつつ (つまり検診方法自体の受診参加率への効果は評価しない)、検診受診者のみを解析した結果に注目する。

医療従事者採取検体による細胞診検査と自己採取検体での hrHPV 検査での比較では、自己採取hrHPV 検査が有意に高い検診パフォーマンスと関

連していた (RR=1.74 [95%CrI: 1.01-2.78])が、結果には大きな異質性が観察され ($\tau=0.40$)、広い予測区間からは細胞診検査が優れる集団の存在が示唆された (図 1.3)。医療従事者採取検体 hrHPV 検査と自己採取検体 hrHPV 検査の比較ではどちらの検体採取法が優れるという証拠は得られなかった (0.90 [95%CrI: 0.54-1.48])。

2. 大腸がん検診

2-1. 前がん病変および大腸がんを検出する感度・特異度

対象者全例に大腸内視鏡を実施し、前がん病変 (advanced neoplasia) および大腸がんを評価した32件を採用とした。7種類の化学法、14種類のヒト免疫法キットデバイスが評価された。ヒト免疫法のカットオフ値は多様であり、化学法と特定のカットオフ値との直接比較は限られていたため、全体で定量的な統合を行った間接比較の結果で評価した。カットオフ別および採取日数別の統合感度・特異度を図2.1に示す。ヒト免疫法を用いた検診はカットオフ値を高くするに従い、統合感度は低下、逆に特異度は上昇するトレード・オフの関係が認められた。便の採取回数およびカットオフ値の変動が10 μ g/gから40 μ g/gの間では統合感度と統合特異度は、比較的安定しており、化学法と比較して感度については優れる傾向であり、特異度については同程度であった。

2-2. 前がん病変および大腸がんを検出する1ラウンド検診パフォーマンスの比較

化学法とヒト免疫法を直接比較し、対象病変 (進行前がん病変および大腸がん) を検出する検診パフォーマンスを評価した無作為化比較試験4件を評価した。子宮頸がんの適用例と同様に、各検診の受診率に関する効果は評価しないため、各便潜血検査を完遂した者 (受診者) に注目した。1ラウンドにおけるパフォーマンスでは化学法と比較して免疫法では前がん病変の検出数は約2.4倍と有意に多かったが (リスク比 [RR] 2.44; 95%CrI: 1.75–3.49; 95%予測区間 [PI]: 1.26–4.84) (図2.2)、がん検出数についてはどちらが有意に優れるという証拠は得られなかった (RR 1.57; 95%CrI: 0.89–2.79; 95%PI: 0.63–3.94) (図2.3)。

D. 考察

子宮頸がん検診で、医療従事者採取検体の有効性評価は確立したが、新しい技術としての自己採取HPV検査の有効性はまだ確立していない。今回、メタアナリシスの手法を用い、2つの検査結果の一致率、感度・特異度、前がん病変であるCIN2+の検出率（1ラウンド検診パフォーマンス）の3つの指標について検討した。

ペアデザインで行われた比較に限定した検査結果の一致率、感度・特異度については、新しい技術と既存の技術はほぼ一致した。また1ラウンド検診パフォーマンスとしての前がん病変の検出率については、HPV検査に関して、新しい技術と既存の技術のどちらが優れるという証拠は得られなかったが、医療従事者細胞診と自己採取HPV検査との比較では、自己採取HPV検査が有意に高い検診パフォーマンスを示していた。この結果の解釈は必ずしも容易ではなく、医療従事者細胞診に比して自己採取HPV検査を強く推奨する結果ではないものの、自己採取HPV検査の効果を否定するものでは少なくともなかった。あくまでCIN2+が治療対象とは限らない（研究が行われた国によって異なる）ためこれだけでは推奨を決める根拠として弱い部分があるものの、代替え指標として検討する意義はあると考えられる。

大腸がん検診については、前がん病変および大腸がんを検出する感度・特異度の便潜血化学法と免疫法の比較で32件の論文が抽出された。このような比較は、WEOなどのガイドラインが代替え指標としての感度・特異度比較を有効性の評価として位置づけたことから、近年非常に多くの研究が報告されていた。これらをメタアナリシスした結果では定量法による免疫法でカットオフ $10\mu\text{g/g}$ から $40\mu\text{g/g}$ の間で統合感度・特異度は安定しており、免疫法が感度で化学法より優れていること、特異度は同等であることが示された。また前がん病変および大腸がんを検出する1ラウンド検診パフォーマンスの比較では4つの無作為化比較試験を評価し、化学法と比して免疫法の前がん病変検出率が約2.4倍高かったが、がんに関してはどちらが優れるという証拠は得られなかった。がんの検出よりも前がん病変の検出を重要視するのであれば、この結果は免疫法を化学法に比して推奨する根拠となりうる。

今回の検討はあくまでエビデンスレポートの一貫として行った中間アウトカム・代替え指標の評価であり、最終アウトカムである死亡率減少効果

がない場合に、がん検診ガイドラインの中でどのように評価するのか、またその条件等は別項に委ねるが、評価の方法としては、わが国でも対応可能であることが示された。

E. 結論

子宮頸がん・大腸がんに関して、中間アウトカムとして感度・特異度、前がん病変・がんの検出率（1ラウンド検診パフォーマンス）の評価をメタアナリシスの手法で検討した。既存の方法と新しい技術の比較という点で評価自体は可能であった。解釈や条件については、別途検討の必要がある。

F. 健康危険情報

特になし

G. 研究発表

1. 論文発表

なし

2. 学会発表

1. 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

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

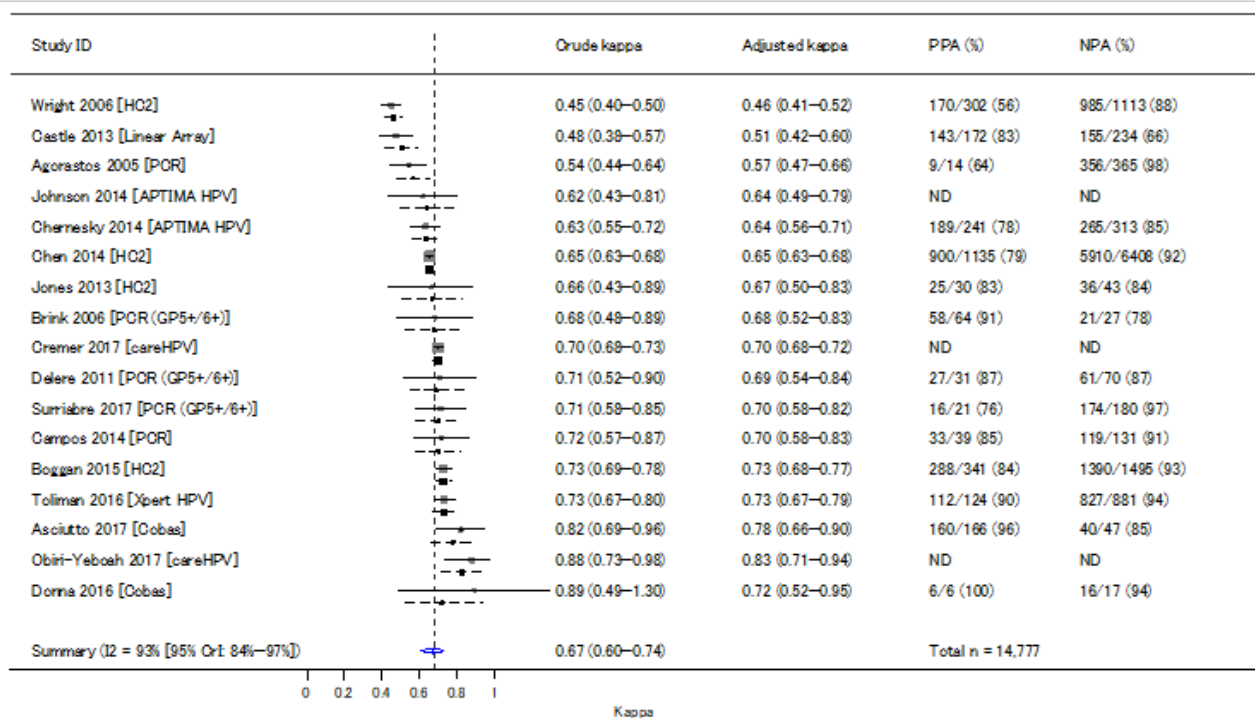


図 1.1. 医療従事者採取検体と自己採取検体での hrHPV 検査の一致度

各研究の補正前 κ 値（実線）、ベイズ階層ランダム効果モデルメタ解析による補正 κ 値（点線）と統合推定値（青色ダイヤモンド）を示す。陽性一致率（PPA）、陰性一致率（NPA）についてはデータ報告があるもののみ示す。ND=データなし。

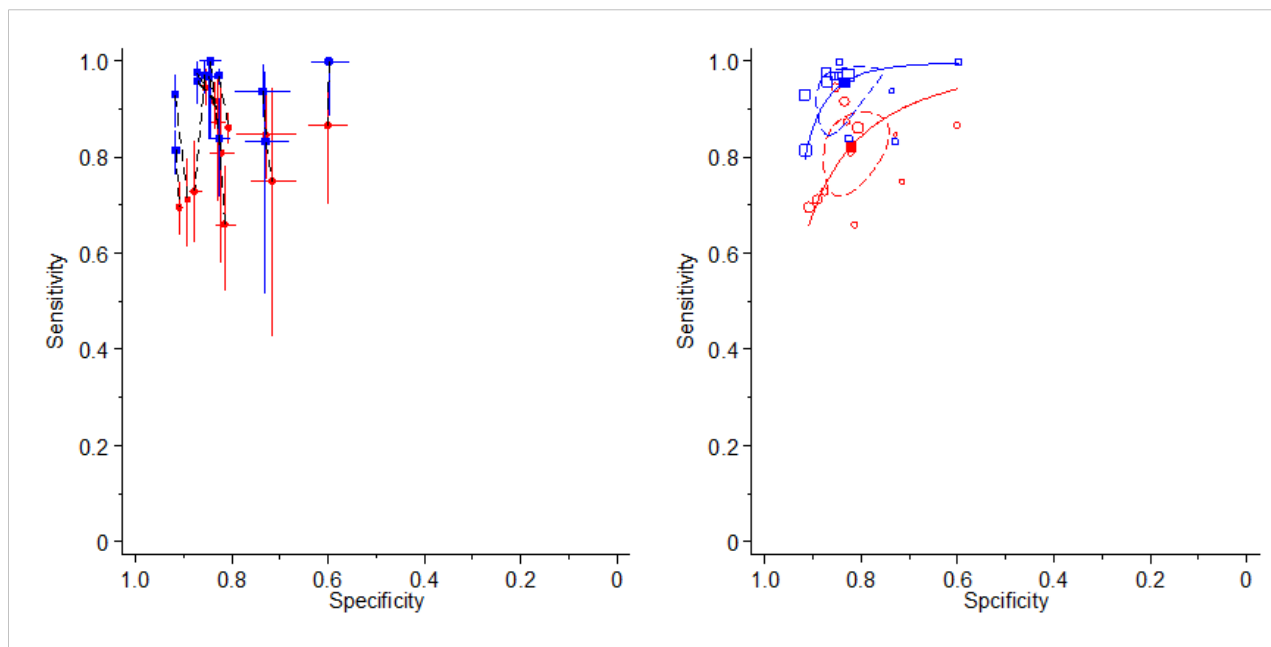


図 1.2. 医療従事者採取検体と自己採取検体での hrHPV 検査の検診精度

医療従事者は青色、自己採取は赤色でそれぞれ表示。報告された補正前精度の十字プロット（左）と各研究の補正精度と統合推定値および信用領域（破線で示す楕円状領域）（右）を示す。

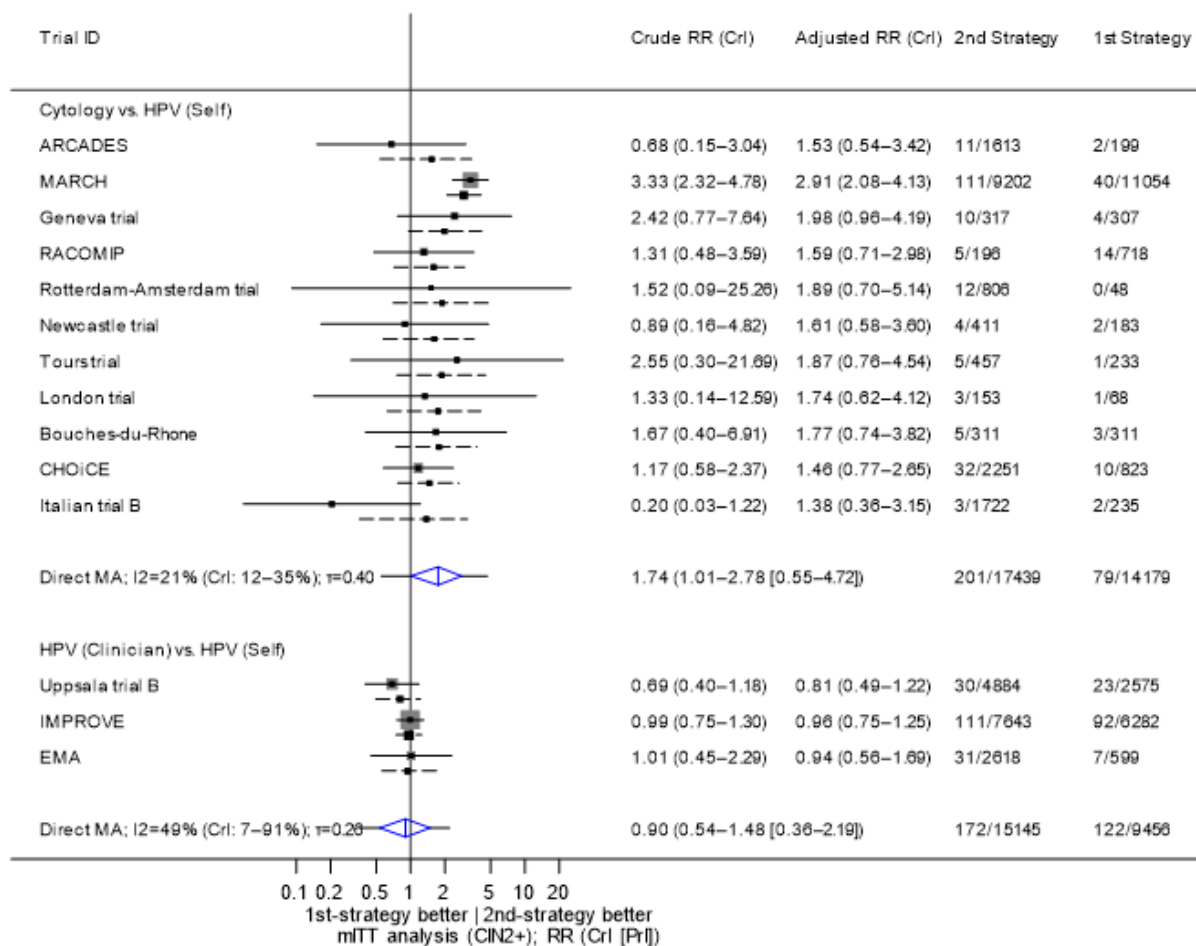


図 1.3. CIN2+を検出する 1 ラウンド検診パフォーマンス

各研究の補正前相対検出率（実線）、ベイズ階層ランダム効果モデルメタ解析による補正相対検出率（点線）と統合推定値（青色ダイヤモンドの幅は 95%ベイズ信用区間、延長する実践は 95%予測区間）を示す。

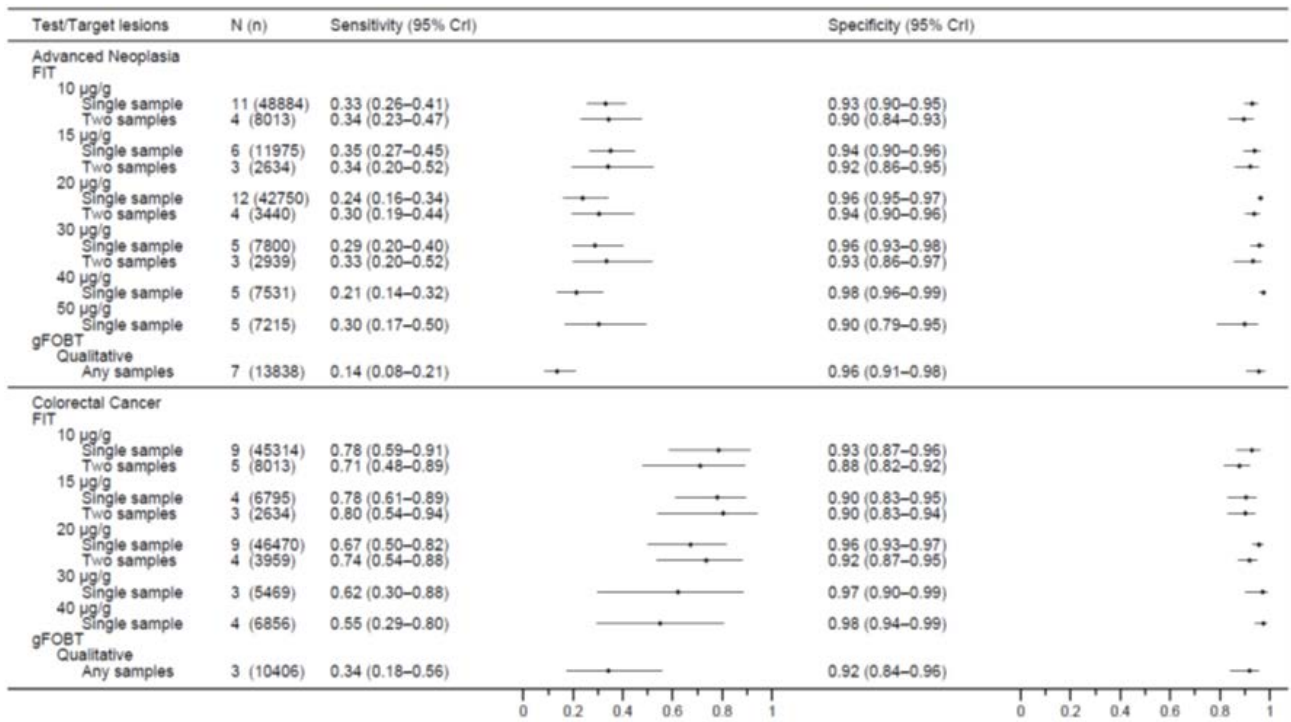


図 2.1. 進行前がん病変および大腸がんを検出する感度・特異度

便の採取回数およびカットオフ値毎に統合感度・特異度および 95% ベイズ信用区間を示す。

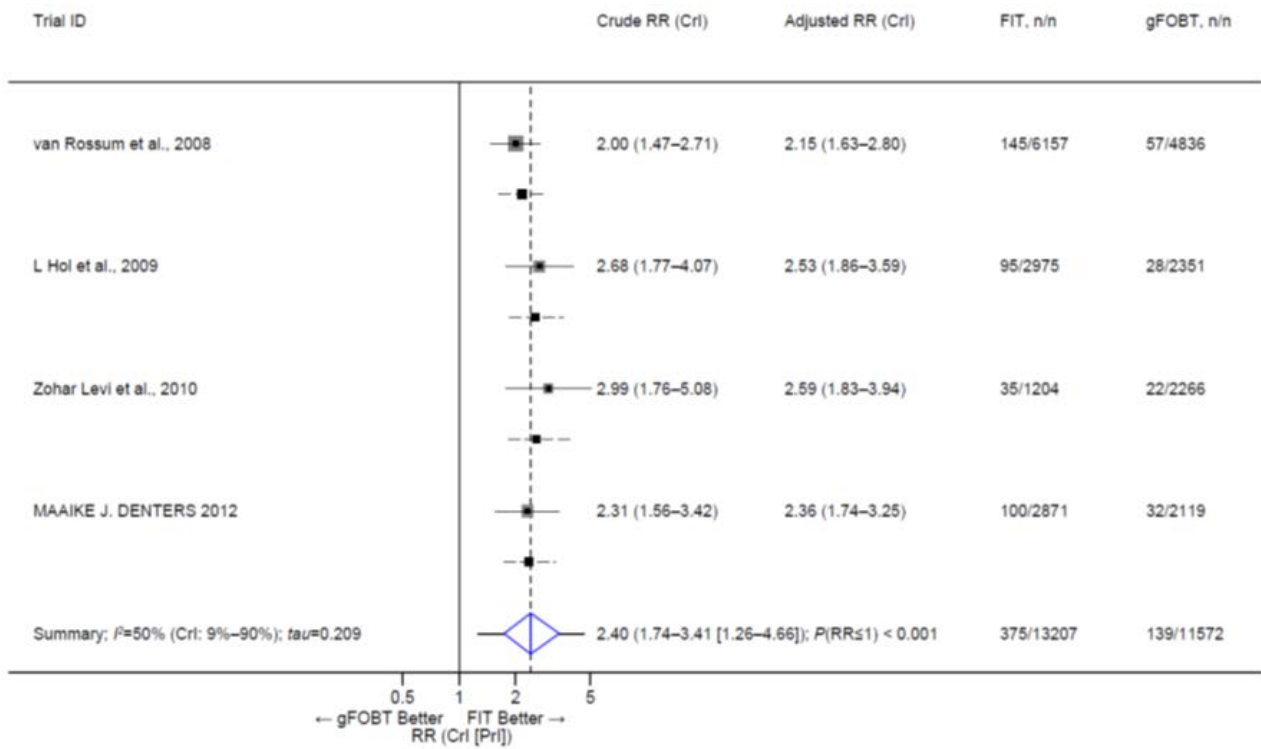


図 2.2. 進行前がん病変を検出する 1 ラウンド検診パフォーマンス

各研究の補正前相対検出率（実線）、ベイズ階層ランダム効果モデルメタ解析による補正相対検出率（点線）と統合推定値（青色ダイヤモンドの幅は 95% ベイズ信用区間、延長する実線は 95% 予測区間）を示す。

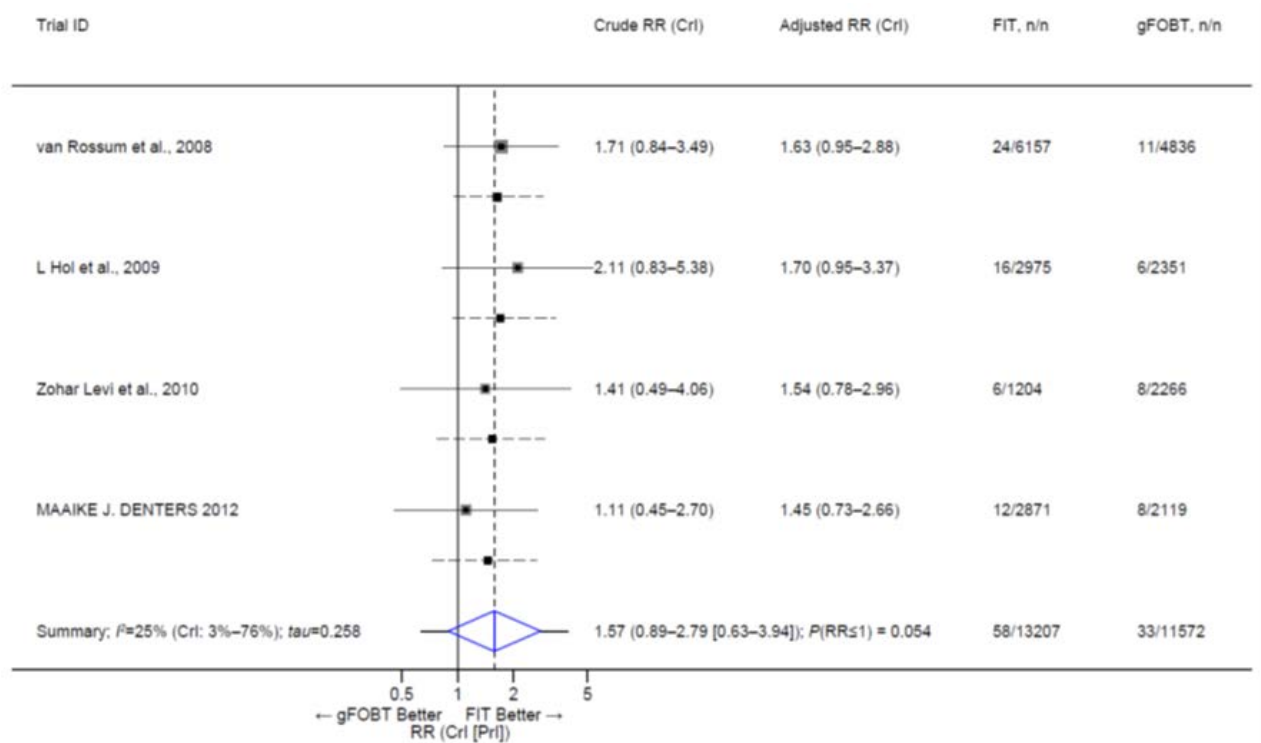


図 2.3. 大腸がんを検出する 1 ラウンド検診パフォーマンス

各研究の補正前相対検出率（実線）、ベイズ階層ランダム効果モデルメタ解析による補正相対検出率（点線）と統合推定値（青色ダイヤモンドの幅は 95% ベイズ信用区間、延長する実線は 95% 予測区間）を示す。