

新しいスクリーニングの効果と効率の評価枠組み
(分担研究：スクリーニングの評価に関する研究)

久繁哲徳, 三笠洋明, 片山貴文

要約

新たなスクリーニングを導入する際の、具体的な評価枠組みを設定することにより、医療政策の意思決定の基礎的な条件を検討したいと考えた。評価枠組みとしては、最も包括的な米国の予防医療特別委員会の枠組みを用い、評価過程における基礎情報の収集と批判的吟味に関する基準を設定した。また、あわせて経済的な効率の評価方法を臨床経済学の枠組みにより設定した。最後に、評価に際しての問題と改善点について考察を加えた。

見出し語：臨床的有効性，経済的効率，検査有効性，効果，研究設計，偏り，偶然，標本数

目的

スクリーニングの導入に際しては、その臨床的効果と経済的効率の評価が前提となる。スクリーニングの効果に関しては、すでにカナダおよび米国の特別委員会（Task Force）によって、包括的な評価基準が提示され、それに基づいて、現行の多様なスクリーニングの評価が実施されている¹⁻³⁾。

ただし、わが国においては、こうした評価が系統的に行われないまま、全国的にスクリーニングが導入されてきた歴史的経過があり、現在においても、実施されているスクリーニングの根拠は必ずしも確立していないのが現状である。しかも、スクリーニングの評価は、全国的に普及した場合、極めて困難な課題となる¹⁻³⁾。

そこで今回は、新たなスクリーニングを導入する際の、具体的な評価枠組みを設定することにより、医療政策の意思決定の基礎的な条件を検討したいと考えた。

評価方法と課題

新しいスクリーニングの評価過程としては、図1および表1に示すように、最も包括的な米国の予防医療特別委員会の枠組みを用いる^{4, 5)}。また、それぞれの過程における評価の基礎情報については、表2に示した方法によって収集を実施する。また、こうして収集した既存情報を系統的・総合的に把握するために、各段階毎に、内容について臨床疫学的な批判的吟味を行う。

第一の〈疾患の負担〉の評価では、スクリーニングの対象となる疾患の発生頻度、その疾患の重症度、予後を評価して、疾患による社会的な負担・損失を総合的に把握、評価する。なお、この評価は、スクリーニングを実施していない場合の疾患の経過・予後、検査、治療法の評価に対応する。

次は、スクリーニングの有効性の評価であるが、この評価には2つの段階がある。一つは、検査の有効性（中間的有効性）、もう一つは健康結果の

徳島大学医学部・衛生

改善（最終的有効性）である¹⁾。

第二の<スクリーニング検査の有効性>の評価では、検査の感度および特異度について評価を行う。とくに、疾患の範囲および偏りの問題については、十分な検討が必要である³⁾。

第三の<スクリーニングの結果>の評価では、健康結果の改善について検討を行う。評価基準としては、国際的にも基準となっている米国のTask Forceの評価基準を用いる^{4, 5)}。基本的には、スクリーニングを実施した場合が、実施しなかった場合に比べて、死亡率・障害の改善が認められるかどうかを評価するが、根拠の質として、研究設計の効力の強さに基づいて順位づけを行う。最も優れた根拠は、無作為臨床試験である。

第四は<スクリーニングの効率>の評価である。スクリーニングの効率については、上記で得られた効果の情報に、費用を組み合わせて評価を行う。ただし、新しいスクリーニングでは、極めて限定された対象・地域でしか実施されていないため、必ずしも正確で十分な情報を把握できない。

最善の方法は、スクリーニングの受診者、非受診者を追跡して、検査、治療の内容と頻度を全て把握し、費用を把握することである。しかし、実際には困難なため、モデルを設定し、シミュレーション的な評価を行う。また、その妥当性をチェックするために症例分析を実施する。また、モデルの設定では、全国的に実施されている標準的な検査・治療の内容について、学会ないし委員会の指針あるいは専門家の意見を参考とする。

最終的に、以上の評価を総合して、スクリーニングの実施可能性について、予測的な検討を実施する。ただし、現時点では、先にも述べたように、効果および効率については、確実な情報を十分に利用できない。したがって、種々の不確実性を考慮に入れて検討した場合、予測的評価の頑健性が保たれるか否かを評価しなければならない。結論が変化する場合には、どのような点に問題があり、それをどの程度改善すべきか、具体的な課題を明らかにすることが求められる。

問題点と考察

上記で今回の評価枠組みを設定したが、スクリーニングの実施可能性を検討する上で、とくに問題となる点について以下に考察を加えてみたい。

1) 検査有効性の評価

検査の有効性で、とくに重要な項目は、感度・特異度の評価である。スクリーニングが普及していない状況および発生率が低い場合は、検査の感度・特異度の評価には注意が必要である。

一つは、<検査対象者の範囲>である。患者については、主要な研究機関に紹介された患者を対象として検査をすると、典型的な患者が選択される可能性が高い。また、対照群では、健康者の中でも比較的健康的な者が選択される可能性が高い。そうすると、感度と特異度が過大に評価される危険性がある。その意味では、対象者の紹介状況と選択の過程に十分な注意が求められる。

次は、<偏り (bias)>である。検査結果の評価に関しては、検閲的検査、検閲的診断、堂々巡りの3種類の偏りが生じる可能性がある^{3, 7)}。その意味では、こうした偏りを防ぐために、ブラインドでの評価が実施されているかどうかを検討することが必要となる。

最後は、<偶然 (chance)>である。報告されている感度および特異度には、偶然による誤差が含まれているため、それらの信頼限界を十分に検討することが求められる^{3, 7)}。例えば、現行の新生児スクリーニングでは、多くの場合感度は90%、特異度は99%を越えている。しかしながら、スクリーニングが広範囲に実施されていない段階で、検査の有効性を検討する場合は、検査対象者の標本数は極めて限られる。

例えば、感度ないし特異度が、90%および95%、99%の場合の、95%信頼区間を図2-図4に示した。信頼区間は、標本数の平方根に反比例する。例えば、90%の場合、標本数が20および50であれば、下方限界は、それぞれ、77%、82%となる。逆に、標本数が5000を越える場合は89%を越える。

したがって、検査対象者数が小例数の場合は、

感度の信頼限界は低くなり、偽陰性率は高くなる。

2) スクリーニングの効果評価

スクリーニングの最終的な有効性である健康改善の評価は、とくに重点をおいて評価することが必要である。この評価において最も重要なのは、繰り返し指摘するように、研究設計である。無作為臨床試験により有効性が示された場合は、スクリーニング実施の強力な根拠となる(表3)。わが国では、スクリーニングの評価において、こうした無作為臨床試験が行われたことは一度も無く、有効性が明確に証明されているプログラムは現在認められない。

無作為臨床試験が困難な場合は、次善の方法としてコホート研究を実施することが望ましい。できるだけ偏りの少ない評価を実施するためには、対象の選択、偏り・攪乱要因の調整など、十分な対応策を視野に入れた研究計画を立てることが求められる。

上記に指摘した系統的誤差以外にも、注意しなければならないのは偶然の誤差である。新生児のマス・スクリーニングの対象疾患は、その頻度が極めて少ないため、偶然による見せかけの効果を慎重に判断する必要がある。

偶然を考慮に入れて、統計学的有意差の評価が可能となるような調査を実施する場合には、比較的大規模な対象数が必要とされる。そこで、有意水準を5%、検出力を80%に設定した際の必要対象数を図5に示した^{7, 8)}(1群当りの数であり、対照群を含めると、実際の対象者数は2倍となる)。

検診により死亡(障害発生)が完全に回避できるような優れたプログラムを想定した場合でも、その率が1/5千では4万人の対象が必要となる。また、1/1万では8万人、1/10万では80万人と急激に増加する。したがって、後者の場合、全国的に調査を実施した場合でも、現在新生児数は年間120万であるため、全員を対象とした追跡調査が必要とされる。

さらに、死亡(障害発生)率が1/100万では、

800万人の対象が必要となり、10年間以上の全新生児を対象とした調査が必要となり、実現はほとんど不可能である。

一方、有効性が上記の半分の検診プログラムを想定した場合は、対象者数はさらに大規模となる。例えば、死亡(障害発生)率が1/1万の場合は、対象者数は50万人におよぶため、この段階で有効性を証明するには、1年間の全新生児を対象とした調査が必要となる。

以上に検討したように、無作為臨床試験ないしコホート研究で評価を行う場合は、対象数も含めた十分な調査計画を組み立てることが必要となる。

しかしながら、死亡(障害発生)率が極めて低い場合には、上記の研究設計による評価は実際には実現不可能である。その場合は、効力の弱い症例-対照研究を実施することになる。ただし、注意しなければならないのは、この方法では、偏りないし攪乱による問題が極めて大きいため、必ずしも明確な根拠が確立できないことである。

3) スクリーニングの効率評価

医療技術の経済的評価には、いくつかの方法があるが(表4)、当初費用-便益分析(cost-benefit analysis)が開発され、その後、広範囲に実施されてきた。しかしながら、生命および健康について金銭的な評価を行うことが条件となるため、代替的な方法として、費用-効果分析(cost-effective analysis)が用いられるようになった。また、近年では、生活の質(quality of life)を考慮に入れた費用-効用分析(cost-utility analysis)が注目されている。

費用-便益分析は、スクリーニングを実施すべきか否か、社会的利益の観点から絶対的な評価を行うことを目的とし、一方、費用-効果、費用-効用分析は、他の比較代替案となる医療技術に比べて、相対的にどの程度効率的に優れているかを評価することを目的としている。したがって、それぞれの限界と利点とを十分に見極めた評価が必要である。

今回も、既存のスクリーニング・プログラムと

同様に、費用-便益分析を実施することを計画している。ただし、上記に挙げた問題点については、さらに検討が必要とされる。その意味では、生活の質を考慮に入れた健康結果に対する金銭評価方法の検討、別の評価方法とくに費用-効用分析の実現可能性の検討など、今後の評価を基礎付ける研究が求められていると考えられる。

文献

1) 久繁哲徳：疾病予防の評価と実行に向けての合意形成, 日本医事新報, 3516:48-52, 1991
 2) 久繁哲徳：マス・スクリーニングのテクノロジー・アセスメント, 効果的なプログラムの選択と優先順位の設定, 日本マス・スクリーニング学会誌, 4:21-29, 1994

3) 久繁哲徳：臨床情報のチェックポイント, ベッドサイドの医療評価学, 医歯薬出版, 1994
 4) Canadian Task Force on the Periodic Health Examination: The periodic health examination, Can Med Assoc J, 121:1193-1254, 1979
 5) U.S. Preventive Task Force: Guide to clinical preventive services, Williams & Wilkins, 1989
 6) 久繁哲徳, 西村周三, 監訳：ドラモンド他, 臨床経済学, 篠原出版, 1990
 7) Sackett DL, et al: Clinical epidemiology, 2nd ed, Little Brown, 1991
 8) 浜島信之：無作為割付臨床試験, 癌と化学療法社, 1993

図1 スクリーニングの評価過程

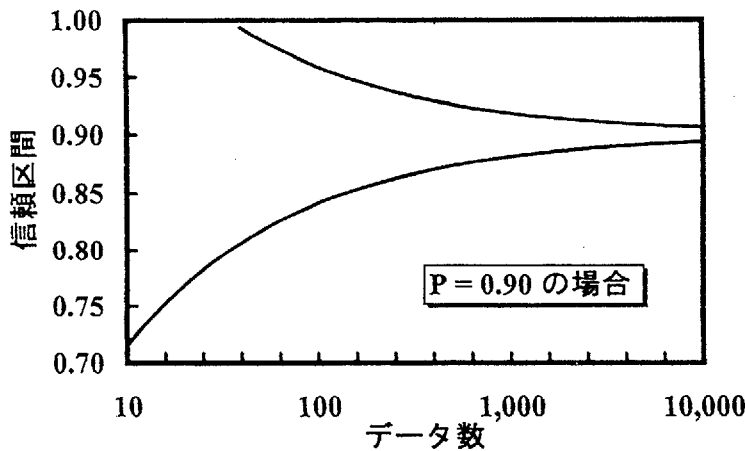
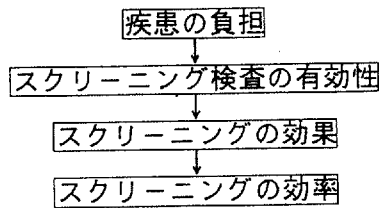


図2 検査の有効性が90%の場合の対象数と信頼区間

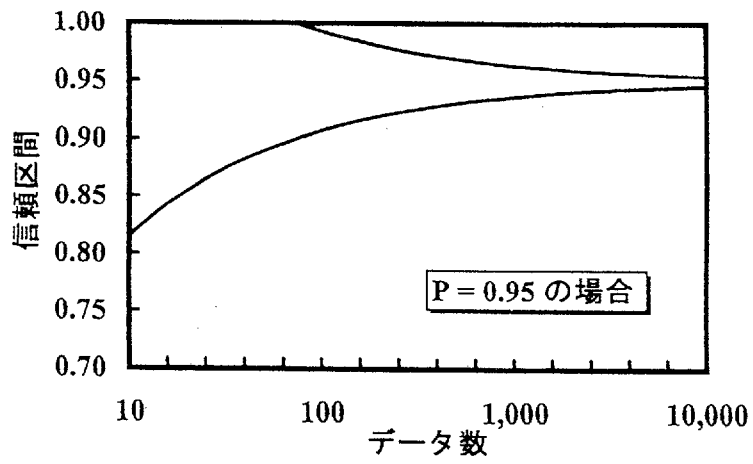


図3 検査の有効性が95%の場合の対象数と信頼区間

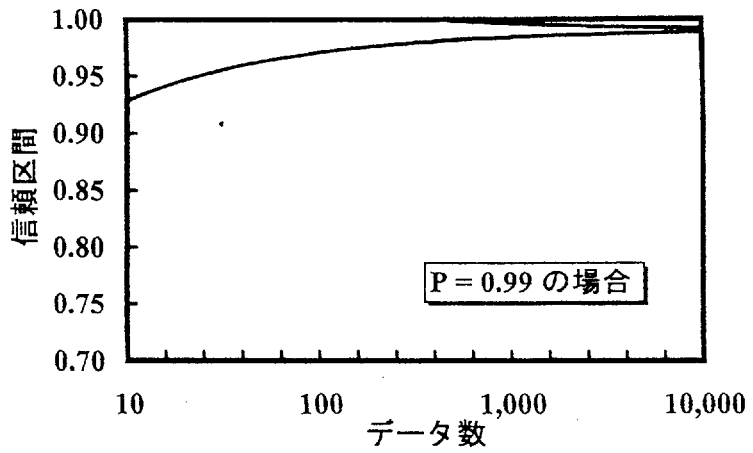


図4 検査の有効性が99%の場合の対象数と信頼区間

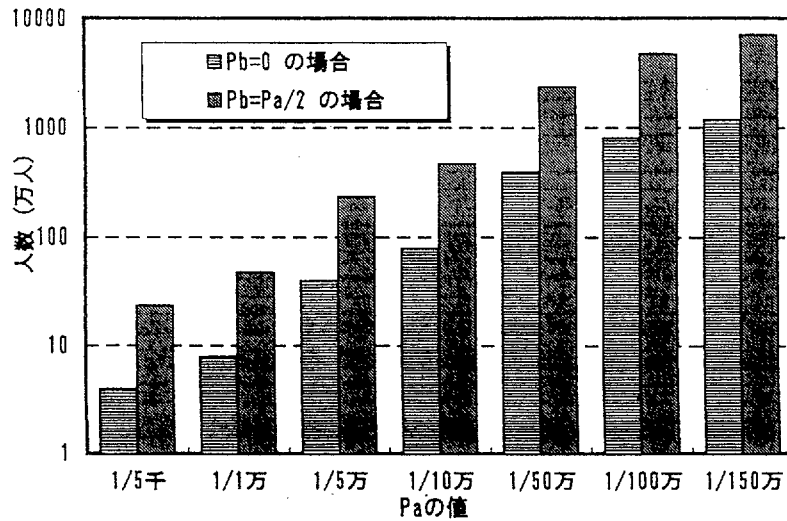


図5 健康改善を評価するための対象数

Pa: スクリーニング非実施群の死亡(障害発生)率
 Pb: スクリーニング実施群の死亡(障害発生)率
 α: 5%, β: 20%

表1 効果評価の基準と勧告の分類(米国予防医学T F)

介入の効果

- I. 適切に設計された無作為臨床試験(RCT)による証拠が少なくとも一つ
- II-1. 優れた設計の無作為化の無い臨床試験(CT)による根拠
- 2. 優れた設計のコホート研究, 症例-対照研究による根拠
(複数の研究機関, 研究集団によることが望ましい)
- 3. 介入の有無を問わない多様な時系列研究による根拠
(意図しない実験による劇的な結果もこうした根拠に含まれる)
- III. 権威者の意見
(臨床経験, 記述研究, 専門委員会の報告に基づく)

勧告の分類

- A. 定期健診の実施を特に推薦すべき条件を満たす優れた根拠がある
- B. 定期健診の実施を特に推薦すべき条件を満たす正当な根拠がある
- C. 定期健診の実施を特に推薦すべき条件を満たす根拠が乏しい
(しかし, 他の根拠から推薦する)
- D. 定期健診の実施を除外すべき条件を満たす正当な根拠がある
- E. 定期健診の実施を除外すべき条件を満たす優れた根拠がある

表2 情報収集の方法

-
- 1) コンピュータ文献検索
 - ・諸外国の論文については, MEDLINEにより検索
 - ・わが国の論文については, JMEDICINEにより検索
 - 2) 関連資料からの文献収集
 - ・書籍, 研究会報告, 会議資料, 総説論文などからの収集
 - ・専門家の研究資料
-

表3 研究設計と効力

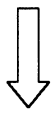
研究設計	効力
無作為臨床試験	最も強い (きっとそうだ)
追跡研究	(そうらしい)
症例-対照研究	(そうかもしれない)
生態学的研究	(そうかな)
症例研究	最も弱い (そうかな)

表4 経済的評価の方法

方法	健康結果の評価
費用-便益分析	金銭
費用-効果分析	身体的単位 (生存年)
費用-効用分析	生活の質 (健康生存年)
費用最小化分析	-



検索用テキスト OCR(光学的文字認識)ソフト使用
論文の一部ですが、認識率の関係で誤字が含まれる場合があります



要約

新たなスクリーニングを導入する際の、具体的な評価枠組みを設定することにより、医療政策の意思決定の基礎的な条件を検討したいと考えた。評価枠組みとしては、最も包括的な米国の予防医療特別委員会の枠組みを用い、評価過程における基礎情報の収集と批判的吟味に関する基準を設定した。また、あわせて経済的な効率の評価方法を臨床経済学の枠組みにより設定した。最後に、評価に際しての問題と改善点について考察を加えた。