があり，通常この問題はRCTでのみ判明する．一方，一般の健康診査で行われている健診項目とそれが対象とする疾患では，必ずしも同様の疾患自然史モデルが適応できない．疾患によっては必ずしもRCTで直接検証しなくとも，その健診項目の効果を考察できる可能性がある．考察で問題となりうるのは，①健康診査の対象疾患がすでにQOLの低下をきたしているのか，②健康診査を受けずイベント発症や症状をきたして受診する場合，これが避けるべき重要なエンドポイントなのか，③健康診査で発見された場合とイベント発症や症状をきたして受診する場合の治療効果の差，などが問題になりうると考えられた．

㉒の2. 2）の「健診項目の効果を検討したほかの観察研究」には症例対照研究やコホート研究が含まれる．また費用効果分析は，既存の報告に基づいたモデルを用いるため多くの仮説設定を要するが，健診の考察に必要な要因を網羅的に検討しているため重要と思われた．

推奨レベルの設定

推奨レベルの設定は，USPSTFの方法に準拠した．このために以下のような方法により，推奨レベルを考察した．

エビデンスの質（quality of evidence）の評価

エビデンスレベルを考察するとき，3つの次元で考察することができる．すなわち，個別の研究ごとのエビデンスレベル，特定のkey questionにおける複数の研究のエビデンスレベル，analytic frameworkの各ステップを通じた連鎖全体でのエビデンスレベルである．ここではエビデンスの質として，RCTなどにより健診効果を直接検討した研究が十分あれば「Good」とした．そのような研究がない場合，analytic frameworkの連鎖がつながる間接的ともいえる一連のエビデンスがあれば「Fair」とした．そしていずれかのステップで研究がされないためエビデンスがなく，連鎖がつながらない場合は「Poor」とした．

㊵推奨レベルの評価

quality of evidence	net benefit			
	Substantial	Moderate	Small	Zero/negative
Good	A	B	C	D
Fair	B	B	C	D
Poor	I	I	I	I

㊶推奨レベル

A	そのような健診項目を実施することが強く勧められる．有効性に関する（対象者の真のアウトカムを改善する）良好なエビデンスがあり，利益は害を非常に上回る．
B	そのような健診項目を実施することが勧められる．有効性に関する（対象者の真のアウトカムを改善する）少なくとも間接的なエビデンスがあり，利益は害を上回る．
C	そのような健診項目を実施することが推奨できるともできないともいえない．有効性に関する（対象者のアウトカムを改善する）少なくとも間接的なエビデンスがあるが，利益は害をわずかに上回るか接近している．
D	そのような健診項目を実施することは推奨できない．無効というエビデンスがあるか，利益より害が大きい．
Ins	そのような健診項目を実施することが推奨できるともできないともいえない．有効性に関する（対象者のアウトカムを改善する）エビデンスはなく，利益と害の比較ができない．

ネットベネフィットの評価

ここでは，有効性と有害性を総合的に考慮したうえで，ネットベネフィットを判定した．しかし主観的になるところであった．

推奨レベルの設定

推奨レベルの決定は，㊵のマトリックスに従いquality of evidenceとネットベネフィットを考慮して，㊶のようにA，B，C，D，Insに設定した．

合意の形成

研究グループ内での合意を形成するため，各担当者が提示した推奨レベルに関して，研究グループ内で検討会を行い，研究グループとして一つの

推奨レベルを提示した.

外部評価

第三者の立場から評価を受けるために，12名の外部評価者に評価を依頼した．依頼した評価者として，プライマリケア，医学教育，予防医学，臨床検査医学，診療ガイドラインに造詣が深い専門家を含めた.

評価のために，研究グループが提示した推奨レベルに対して外部評価者がどの程度同意するかを9段階で示してもらった．1〜3は同意できない場合，4〜6は同意できるともできないともいえない場合，7〜9が同意できる場合に相当する．平成17年度の時点では，評価結果に基づいた推奨レベルの変更は行わず，評価結果をそのまま記載することとした.

結果

健診項目とその対象疾患，そして研究グループが提示する推奨レベルは⑯のようであった．各健診項目に関するエビデンスの要約についてはここでは割愛する.

全体を通じて顕著なことは，健診項目の有効性を直接検討したRCTや観察研究がほとんどないことであった．これは癌検診でこのような研究が盛んに行われてきたことと大きく異なっている．外部評価の結果では全24項目のうち19項目では同意スコア中央値は7以上の同意レベルであった．同意スコア中央値が7未満であったのは，6.5が肝機能と心電図，5.5であったのが尿蛋白，血液一般，胸部X線であった.

外部評価者の総合的コメントではいくつか重要な指摘があり，これらを要約すると次のようである.

① 健康診査の評価として，直接的なRCTをどの程度重要視するべきかは大きな問題である.
② 個別の健診項目と個別の対象疾患という枠組みでの分析検討では不十分な可能性がある.
③ 癌検診との違いを考慮しつつ，潜在疾患の発見という目的だけでなくリスク評価という目的での有効性評価を考慮すべきである.
④ 疾患により健康診査の役割に差がある．analytic frameworkとkey questionを重視すべきである.
⑤ 各健診項目により推奨レベル，特にInsの扱いが異なっている.
⑥ 推奨レベルInsが「無効」であることと等価ではないことを強調すべきではないか.

今後

平成17年度までで健診項目の系統的評価を行い，推奨レベルを提示した．一部の健診項目に関しては，外部評価者の同意レベルは高くなく，これらの健診項目の記載や推奨レベルについてはさらに吟味が必要である．また健診項目ごとに専門診療科，公衆衛生専門家，地域保健に携わる医療従事者，など幅広い分野からの参加者やstakeholderも交えた作成パネルを形成し，十分エビデンスを検証したうえで意見を集約する必要がある．そして複数の段階からなる外部評価を十分受けつつ改善をはかる必要がある．さらにいくつかの健診項目については，なお検討できていない．たとえば尿潜血，骨密度測定，脳MRIなどは近年普及しており，これらに関する評価も必要となる.

このような健診項目の評価は継続的に必要であり，多方面のエビデンスの収集，データベースの構築，作成体制・評価体制の構築，記載の更新，情報の公開，などを行い随時社会の要請に応えられることが望ましい．わが国では海外と比べて健診や人間ドックの普及，画像診断の普及といった特殊な事情がある．このようななかで，独自のエビデンスの創出やガイドライン作成が求められている.

近年医療情報が一般のなかにも普及しているなかで，社会一般の健康診査に対する関心も高い．一般の正しい理解を促すような正確で有用な

⑯健診項目の推奨レベル

健診項目	対象疾患	推奨レベル
喫煙についての問診	喫煙	A
飲酒に関した問診	問題飲酒	B
うつに関する問診	うつ状態	B
自殺に関する問診	自殺高リスク状態	Ins
身長，体重(BMI)	肥満	B
血圧	高血圧	A
身体診察	非特定	Ins
聴力に関する問診(高齢者)	聴力障害	B
聴力検査(一般人)	聴力障害	D
聴力検査(高齢者)	聴力障害	C
視力検査(一般人)	視力低下	Ins
視力検査(高齢者)	視力低下	B
MMSEなど認知症問診	認知症	Ins
検尿(尿蛋白)	蛋白尿	Ins
検尿(尿糖)	糖尿病	D
血液一般	鉄欠乏性貧血	Ins
空腹時血糖，ブドウ糖負荷試験，HbA_{1c}	糖尿病	B
脂質	高脂血症	B
尿酸	高尿酸血症	Ins
肝機能	脂肪肝	Ins
HBV抗原	HBVキャリア・B型慢性肝炎	C
HCV抗体	HCVキャリア・C型慢性肝炎	C
心電図，負荷心電図	虚血性心疾患	Ins
胸部X線	肺癌	Ins

*：胸部X線については，「EBMの手法による肺癌診療ガイドライン2003年版」では胸部X線検査と喀痰細胞診を用いた肺癌検診の推奨グレードをBと記載している[11]．

情報の提示が求められている．しかし，このような健診項目に対する系統的評価の試みは国内では端緒についたばかりである．評価や情報の提供がより充実し，適切な医療技術が正しく理解されたうえで利用されることが望まれる．

◎文献

1. 健康診査に関する各制度の比較
   http://www.mhlw.go.jp/shingi/2005/07/s0725-7b01.html
2. http://www.ahrq.gov/clinic/uspstfix.htm
3. 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(班長祖父江友孝)．有効性評価に基づく大腸がん検診ガイドライン．
4. 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(班長祖父江友孝)．有効性評価に基づく胃がん検診ガイドライン．
5. 厚生労働省科学研究費「疾病予防サービスの制度に関する研究班」(永井良三班長)報告平成18年「わが国における基本的健康診査の検診項目のエビデンスに関する研究」(分担研究者 福井次矢)
6. 福井次矢，上塚芳郎，内山 伸ほか．検診項目評価要約版．印刷中．

7. 厚生労働科学研究費補助金「最新の科学的知見に基づいた保健事業に係る調査研究」(主任研究者 福井次矢) 平成16年度総括・分担研究報告書.
8. 祖父江友孝, 酒井ちさと, 廣瀬 昶ほか. 有効性評価に基づくガイドライン作成手順 (普及版). 癌と化学療法 2005 ; 32 : 893-900.
9. 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). 有効性に基づくがん検診ガイドライン作成手順. 2005.
10. 矢野栄二. 健康診断の有効性と有用性. 産業医学レビュー 2002 ; 1 : 1-31.
11. Evidence-based Medicine (EBM) の手法による肺癌の診療ガイドライン策定に関する研究班. EBMの手法による肺癌診療ガイドライン2003年版. http://minds.jcqhc.or.jp/to/index.aspx

総説

健診項目のエビデンス

福井次矢[1)]

〔SUMMARY〕 2004・05(平成16・17)年度の厚生労働科学研究で，一般健診項目の有効性評価を行った．エビデンス・レベルを評価した後，健診項目としてどの程度奨められるのか評価した．ランダム化比較試験などで健康アウトカムが改善するかどうか直接検証した研究は少なく，様々な研究の結果を組み合わせて(研究の連鎖で)判断が必要な項目が多かった．

メタボリックシンドローム関係の血圧測定(推奨レベルA)，身長と体重(BMI)の測定，空腹時血糖・HbA1c，脂質の測定(推奨レベルB)は，健診項目の要件をほぼ満たしていた．〔臨床検査　51：1181-1185，2007〕

〔KEYWORDS〕 健診項目，エビデンス・レベル，推奨レベル

はじめに

なんらかの症候を訴えて医療機関を訪れる患者を対象に行われる検査は，それが検出しようとする異常状態・疾患を正しく識別でき，患者には危害を与えず，検査を行う側にとっては簡便で，しかもコストが手ごろな範囲内に収まることが求められる．健康診査(健診)では，なんら症候のない(一見)健康な人々を対象とするため，上記の条件がいっそう厳しいものになる．つまり，異常状態・疾患をほぼ見逃すことなく(感度が高い)，健康な人を対象とすることから危害を与える可能性はゼロに近く，膨大な数の受診者を扱うため簡便性はより高く，そしてコストもできる限り安価でなくてはならない．

健診で行う価値のある検査とは，さらにいくつかの条件をクリアするものでなくてはならない．検出しようとする異常状態・疾患は放置し自然経過に任せると重大なアウトカムになりうること，その発症率・罹病率が高いこと，確定診断のために必要となる二次("精密")検査の負担(危害，コストなど)があまりにも大きくないこと，そもそも，発見された異常状態・疾患に対して有効な治療法があり，症候が現れる前に治療(早期治療)すれば，症候が出現してから治療するよりもよい結果になること，などである．

このような条件をすべて満たす健診項目はいったいどのようなものなのかについては，北米ではすでに1970年代後半から科学的な方法で評価され，その結果が速やかに厚生行政に反映されてきた．一方，わが国では，癌検診については1990年代後半になって詳細な評価〔厚生労働科学研究費補助金「がん検診の有効性評価に関する研究」(主任研究者：久道茂)平成10年度報告書〕が行われるようになったが，一般健診項目について体系的に評価したのは，2004年度にわれわれが着手した研究〔厚生労働科学研究費補助金「最新の科学的知見に基づいた保健事業に係る調査研究」(主任研究者：福井次矢)平成16年度総括・分担研究報告書〕が最初のもののようである．

本稿では，上記のわれわれの研究とそれを要約した報告書〔厚生労働科学研究費補助金「最新の科学的知見に基づいた保健事業に係る調査研究」基本的健康診査の検診項目のエビデンスに基づく

1) FUKUI Tsuguya　聖路加国際病院・院長

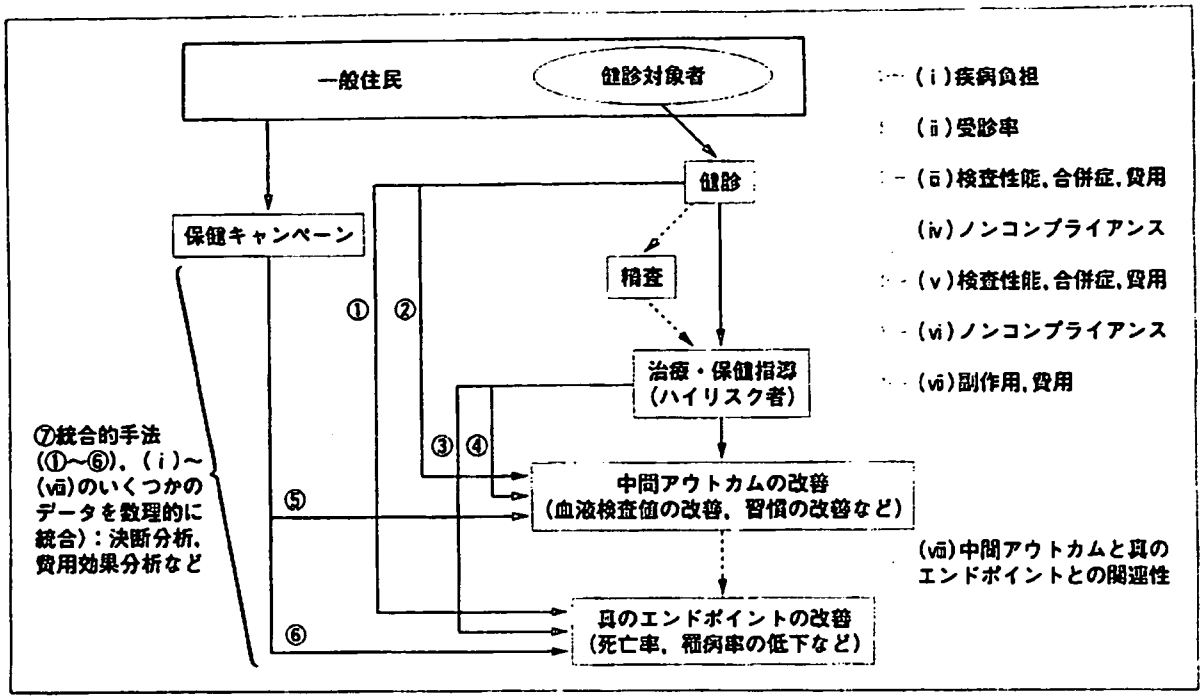

図1　健診・保健事業評価モデル

表1　エビデンス・レベル

Level 1	ランダム化比較試験
Level 2	非ランダム化比較試験
Level 3	コホート研究
Level 4	症例対照研究
Level 5	症例シリーズ
Level 6	専門家の主観的意見

評価に係わる研究(分担研究者：福井次矢)平成17年度分担研究報告書〕で採用した評価方法，評価結果の一部を紹介する．

健診項目のエビデンスの有無とレベル

2004(平成16)年度の研究では，まず，各研究者の担当項目を決め，EBM (evidence based medicine)の手順で文献検索を行った．そして，比較的レベルの高い文献を選択してエビデンステーブルを作成，そのうえで，検診項目のアウトカムに及ぼす効果についてどのレベルのエビデンスがあるのか評価した．

健診の項目が健康アウトカムを改善するかどうかを直接ランダム化比較試験を行って検証している研究は少ないため，様々な研究の結果を組み合わせて，最終的にアウトカムを改善するかどうか判断せざるを得ない．したがって，健診項目とアウトカムをつなぐどの段階についての研究なのかを明示するために，健診・保健事業評価モデル(図1)を作成した．

中間エンドポイント(intermediate endpoint：IE，血液検査値の改善，習慣の改善など)，真のエンドポイント(true endpoint：TE，死亡率の低下や罹病率の低下など)のどちらについて評価がなされているのかをも明示した．

エビデンス・レベルは6段階で表示した(表1)．

健診項目と検出対象疾患，評価結果(エンドポイントの種類とエビデンスレベル)を表2に示す．

総体的な結果としては，表3のようになり，専門家の意見さえ記載されている文献が見つからない健診項目が6を数えた．

健診項目の推奨レベル

2005年度の研究では，前年度のデータについて，USPSTF (U.S. Preventive Services Task Force)の方法を応用して，健診項目として推奨できるかどうか，推奨レベルを決定した．

2004年度の研究結果を基に，エビデンスのレベルを個別の研究ごとに考えるだけでなく，健診

表2　評価結果

1. 飲酒に関する問診	問題飲酒	TE	Level 1
2. 喫煙に関する問診	喫煙	TE	Level 1
3. うつに関する問診	うつ状態	TE	Level 1
4. 自殺に関する問診	自殺高リスク状態	—	Level 6
5. 認知症に関する問診	認知症	TE	Level 1
6. 身長と体重(BMI)の測定	肥満	TE	Level 1
7. 血圧測定	高血圧	TE	Level 1
8. 視力測定	視力低下	—	Level 6
9. 聴力測定	聴力障害	—	Level 6
10. 身体診察	非特定	—	Level 6
11. 聴診	非特定	エビデンスは見つからなかった	
12. 腹部の診察	非特定	エビデンスは見つからなかった	
13. 安静時12誘導心電図および運動負荷心電図	虚血性心疾患	TE	Level 1
14. 胸部X線写真	肺癌	TE	Level 1
15. 呼吸機能検査	非特定	エビデンスは見つからなかった	
16. 空腹時血糖，糖負荷試験	糖尿病	TE	Level 1
17. 尿酸	高尿酸血症	エビデンスは見つからなかった	
18. HBV抗原	HBVキャリア・B型慢性肝炎（予防接種の有効性を示す費用効果分析あり）	TE	Level 3
19. HCV抗体	HCVキャリア・C型慢性肝炎	エビデンスは見つからなかった（予防接種の有効性を示す費用効果分析あり）	
20. 血清コレステロール	動脈硬化性疾患	IE	Level 1
21. 中性脂肪	動脈硬化性疾患		Level 6
22. AST・ALT・γ-GTP	脂肪肝	IE	Level 5
23. 尿蛋白	腎疾患	エビデンスは見つからなかった（統合型研究あるも結論は一定していない）	
24. 尿糖	糖尿病		Level 6
25. 血球検査	鉄欠乏性貧血	TE	Level 1

表3　健診25項目のエビデンス(まとめ)

レベル1の検診項目	11項目
・効果あり	9項目
・効果なし	2項目
レベル3	1項目
レベル5	1項目
レベル6	6項目
エビデンスなし	6項目

項目とアウトカムをつなぐどの部分のエビデンスがあるのか，欠けているのかなどを考慮して連鎖全体でのエビデンスを評価した．健診項目評価の分析的枠組みを用いて，どの課題についての答えを得ようとしている研究なのかを明確にした(図2)．健診項目のアウトカムを直接検討した研究があれば「Good」とし，そのような研究はないが複数の研究をつなぎ合わせればアウトカムに到達するときには「Fair」，そして連鎖が繋がるようなエビデンスがない場合には「Poor」とした．

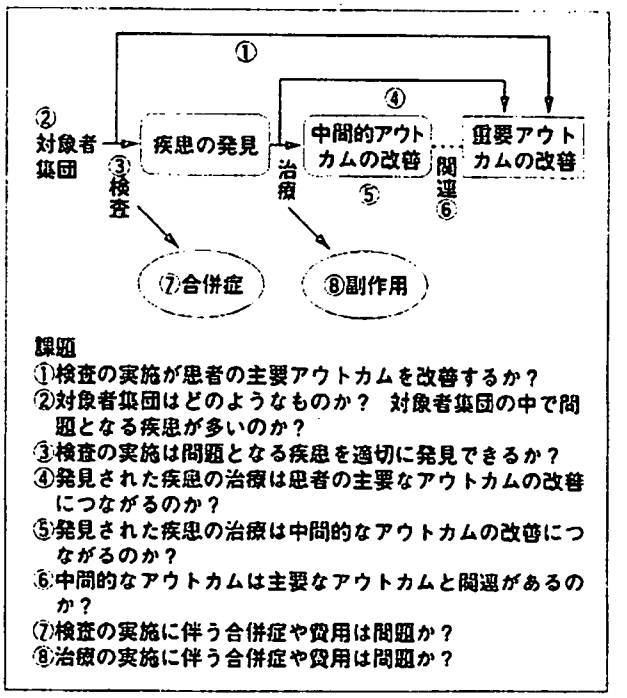

図2　健診項目評価の分析的枠組と課題

表4　推奨レベルの分類

A＝当該検査を健診で行うことが強く奨められる．アウトカムを改善するというレベルの高いエビデンスがあり，利益は害を大きく上回る．
B＝当該検査を健診で行うことが奨められる．アウトカムが改善するという少なくとも間接的なエビデンスがあり，利益が害を上回る．
C＝当該検査を健診で行うよう奨めるとも奨めないとも言えない．間接的なエビデンスはあるが，利益は害をわずかに上回るか接近している．
D＝当該検査を健診で行うことは奨められない．当該検査が健診項目としては無効とのエビデンスがあるか，あるいは利益より害が大きい．
Ins＝当該検査を健診項目として推奨することの判断ができない．アウトカムを改善するというエビデンスがなく，利益と害の比較ができない．

表5　推奨レベルのマトリックス

エビデンスのレベル	正味利得(Net Benefit)			
	大	中等度	小	ゼロ/害
Good	A	B	C	D
Fair	B	B	C	D
Poor	Ins	Ins	Ins	Ins

表6　評価結果

1. 飲酒に関する問診	問題飲酒	B
2. 喫煙に関する問診	喫煙	A
3. うつに関する問診	うつ状態	B
4. 自殺に関する問診	自殺高リスク状態	Ins
5. 認知症に関する問診	認知症	Ins
6. 身長と体重(BMI)の測定	肥満	B
7. 血圧測定	高血圧	A
8. 視力測定(一般成人)	視力低下	Ins
9. 視力検査(高齢者)	視力低下	B
10. 聴力に関する問診	聴力障害	B
11. 聴力検査(一般成人)	聴力障害	D
12. 聴力検査(高齢者)	聴力障害	C
13. 身体診察	非特定	Ins
14. 安静時12誘導心電図および運動負荷心電図	虚血性心疾患	B
15. 胸部X線写真	肺癌	Ins
16. 空腹時血糖，糖負荷試験，HbA1c	糖尿病	B
17. 尿酸	高尿酸血症	Ins
18. HBV抗原	HBVキャリア・B型慢性肝炎	C
19. HCV抗体	HCVキャリア・C型慢性肝炎	C
20. 脂質	高脂血症	B
21. AST・ALT・γ-GTP	脂肪肝	Ins
22. 尿蛋白	腎疾患	Ins
23. 尿糖	糖尿病	D
24. 血球検査	鉄欠乏性貧血	Ins

健診による有効性と有害性を総合的に考えて，主観的にはなるが，正味利得(Net Benefit)を判断した．

推奨レベルを，表4のような5段階に分類した．

したがって，推奨のレベルを，研究のエビデンス・レベルと正味利得の2次元マトリックスに表すと表5のようになる．

以上の方法によって評価した結果を表6に示す．

健診項目の有効性評価

以上より，一般健診で行われている検査の有効性を直接評価した研究は少なく，ましてやランダム化比較試験や質の高い観察研究はほとんど行われてきていないことが明らかとなった．これは癌検診で質の高い研究が多く行われてきていること

と大きく異なっている．有効であるに違いないという信念が広く根強く，すでに当然のように行われている現状を考えると，ランダム化比較試験のように科学的に厳密な方法論での評価は困難であろう．しかし，社会全体としては莫大なコストがかかっていることを考えると，少なくとも科学的に納得の行くレベルのエビデンスを得るための評価研究は必要であろう．

メタボリックシンドローム関係の健診項目

血圧測定については，アウトカムの改善が実証されていて，検査自体に害を及ぼす可能性もなく，推奨レベルAである．身長と体重(BMI：body mass index)の測定，空腹時血糖やHbA1c，脂質の測定などは推奨レベルBであり，現在のところ，健診項目としての要件はほぼ満たしていると考えてよい．

ELSEVIER

International Journal of Cardiology 118 (2007) 88–96

International Journal of Cardiology

www.elsevier.com/locate/ijcard

Cost-effectiveness of coronary artery disease screening in asymptomatic patients with type 2 diabetes and other atherogenic risk factors in Japan: Factors influencing on international application of evidence-based guidelines

Yasuaki Hayashino [a,*], Takuro Shimbo [b], Satoru Tsujii [c], Hitoshi Ishii [d], Hirokazu Kondo [e], Tsukasa Nakamura [f], Shizuko Nagata-Kobayashi [g], Tsuguya Fukui [h]

[a] *Department of Epidemiology and Healthcare Research, Kyoto University Graduate School of Medicine, Konoe-cho, Yoshida, Sakyo-ku, Kyoto 606-8501 Japan*
[b] *Department of Medical Ecology and Informatics, Research Institute, International Medical Center of Japan, Tokyo, Japan*
[c] *Tenri Hospital Diabetes Center, Nara, Japan*
[d] *Department of Endocrinology, Tenri Hospital, Nara, Japan*
[e] *Department of Geriatric Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan*
[f] *Department of General Medicine, Shimane Prefectural Central Hospital, Izumo, Japan*
[g] *Department of General Medicine and Clinical Epidemiology, Kyoto University Graduate School of Medicine, Kyoto, Japan*
[h] *St. Luke's International Hospital, Tokyo, Japan*

Received 24 March 2005; received in revised form 9 February 2006; accepted 24 March 2006
Available online 1 September 2006

Abstract

Background: Screening for coronary artery disease (CAD) in asymptomatic diabetic patients with atherogenic risk factors is recommended by the American College of Cardiology/American Diabetes Association. It is not clear whether these guidelines apply to the Japanese population with a different epidemiology of CAD. This study evaluates the applicability of the U.S. guidelines to Japan, taking account of cost-effectiveness.

Design: A cost-effectiveness analysis using a Markov model was performed to measure the clinical benefit and cost of CAD screening in asymptomatic patients with diabetes and additional atherogenic risk factors. We evaluated cohorts of patients stratified by age, gender, and atherogenic risks. The incremental cost-effectiveness of not screening, exercise electrocardiography, exercise echocardiography, and exercise single-photon emission-tomography (SPECT) was calculated. The data used were obtained from the literature. Outcomes are expressed as US dollars per quality-adjusted life year (QALY).

Results: Compared with not screening, the incremental cost-effectiveness ratio (ICER) of exercise electrocardiography was $31,400/QALY for 60-year-old asymptomatic diabetic men, and 46,600 for 65-year-old women with hypertension and smoking. The ICER of exercise echocardiography was $31,500/QALY and of SPECT was $326,000/QALY, compared with the next dominant strategy. Sensitivity analyses found that these results varied according to age, gender, the combination of additional atherogenic risk factors, and the frequency of screening.

Conclusion: From a societal perspective the U.S. guidelines on screening for CAD in high risk diabetic patients are applicable to the Japanese population. However, the population subjected to screening should be carefully selected to obtain greatest benefit from screening.

Keywords: Silent myocardial ischemia; Cost-effectiveness analysis; Screening strategy; Diabetes mellitus; International comparison

1. Introduction

Coronary artery disease (CAD) is not a major cause of morbidity and mortality in the Japanese population. The

* Corresponding author. Tel.: +81 75 753 9467; fax: +81 75 753 4644.
E-mail address: hayasino-y@umin.net (Y. Hayashino).

0167-5273/$ - see front matter © 2006 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.ijcard.2006.03.086

prevalence of CAD estimated from survey data is 6.3 per 1000 persons in Japan (1999) [1], which is much less than the U.S. prevalence of 46 per 1000 persons (2004) [2]. However, in diabetic patients, CAD is a major problem even in Japan. The 2002 Japanese survey on diabetic patients recently released by the Ministry of Health, Labor and Welfare estimated that the prevalence of CAD in diabetic patients was 158 per 1000 persons (2002) [3], which is much higher than in the general population. In diabetic patients, CAD is usually diagnosed at an advanced stage and has a correspondingly dismal prognosis [4,5]. The delay in diagnosing CAD is due partly to the presence of asymptomatic myocardial ischemia, thus it is very important to identify CAD in patients with asymptomatic diabetes.

Currently the Japanese Circulation Society/Japanese Diabetes Society does not have explicit criteria on CAD screening, while the American College of Cardiology/American Diabetes Association recommends that cardiac testing be done irrespective of the presence of CAD symptoms in diabetic patients with two or more atherogenic risk factors, due to the high prevalence of CAD in diabetic patients [6]. However, these guidelines cannot simply be applied to Japanese patients, because the incidence and prevalence of CAD are different in these two countries. A cost-effectiveness analysis may be a useful method to evaluate the applicability of guidelines to different countries with different characteristics [7]. In a previously published study, we have shown that screening for CAD is cost-effective in high risk diabetic patients in the U.S. [8]. The same model with Japanese data could also be used to compare the cost-effectiveness of CAD screening in these two countries.

To determine whether U.S. guidelines are applicable to Japan's situation with regard to cost-effectiveness ratio, we used a Markov model to perform a cost-effectiveness analysis of different screening strategies. We then examined the costs and benefits of CAD screening for asymptomatic diabetic patients with two additional atherogenic risk factors in the Japanese population.

2. Methods

We evaluated cohorts of patients stratified by age (55, 60, 65, 70 years of age), and three pairs of atherogenic risk factors (hypertension, smoking, and low-density lipoprotein (LDL) level > 160 mg/dl), as recommended by the American College of Cardiology/American Diabetes Association [6]. We excluded the high-density lipoprotein (HDL) level as a further risk factor because we could not find enough data to estimate its influence on the prevalence of CAD in Japan. The base case cohort consisted of asymptomatic men with type 2 diabetes mellitus and two additional atherogenic risk factors. Based on Cohn's classification [9], we defined silent myocardial ischemia as asymptomatic ischemia without a history of angina/myocardial infarction (MI). Since most available data relates to hypertensive men who are approximately 60 years old and smoke, the base-case analysis was conducted on this group. A Markov model was used to estimate the lifetime costs and quality-adjusted life years [10].

We used the DATA 3.5.9 software (TreeAge Software, Inc, Williamstown, MA, USA) to calculate costs and outcomes. Costs were estimated from a societal perspective, and outcomes were measured in quality-adjusted life years (QALYs). We then calculated the incremental cost-effectiveness for all competing strategies.

2.1. Decision-analytic model

To compare the situations in the U.S. and Japan, we used the model that we used to evaluate the cost-effectiveness of screening in the U.S. [8], but with Japanese data. Here, we

Table 1
Baseline values and ranges in sensitivity analysis

Variables	Baseline value	Lower range	Upper range	Reference
Prevalence of asymptomatic ischemia in base-case (estimated)[a]	0.32	0.22	0.42	[15,16,26,63]
Annual incidence of CAD in base-case, per 1000 (estimated)[a]	13.7	9.6	17.8	[13,14]
Proportion of silent ischemia in patients with myocardial ischemia[b]	0.4	0.38	0.62	[18]
Costs ($2003)[a]				
Exercise electrocardiography	54	38	70	
Exercise echocardiography	104	73	135	
Exercise SPECT	731	511	950	
Coronary angiography	578	405	751	
PTCA	18713	13099	24327	
CABG	32685	22879	42490	
Annual cost				
Symptomatic myocardial ischemia	1845	1291	2398	
History of MI	1766	1236	2296	
Conventional diabetes care	1934	1354	2514	
Patient time cost	2169	1518	2819	[1,44]
Aspirin	18	12	23	
One time cost[a]				
MI death	20002	14002	26003	
MI survive	16351	11446	21257	
Health utility (TTO)[a]				
Symptomatic myocardial ischemia	0.73	0.62	0.84	
History of MI	0.60	0.44	0.74	

SPECT indicates single-photon emission-tomography; CAD, coronary artery disease; MI, myocardial infarction; PTCA, percutaneous transluminal coronary angioplasty; CABG, coronary artery bypass grafting.
[a] Ranges based on +/−30% of base-case estimates.
[b] Ranges based on reported proportions in literature.

provide only an outline of the model. Table 1 shows the estimates used in the current analysis including health utilities and cost data which are different from our U.S. analysis. Our model consisted of seven CAD states: normal, silent ischemia, symptomatic ischemia, history of myocardial infarction, post-percutaneous transluminal coronary angioplasty (PTCA), post-coronary artery bypass grafting (CABG), and death [8,11]. The prevalence of CAD for each risk state was calculated from the prevalence of CAD in the general Japanese population and the relative risk associated with atherogenic CAD risk factors stratified by age and gender. The baseline prevalence of CAD in different age and gender groups was derived from patient survey data provided by the Ministry of Health, Labor and Welfare in Japan [1]. The relative risks for the Japanese population were taken from the published literature [12].

Transition probabilities of moving from the normal state to the CAD state were determined according to demographic and epidemiological factors. These include age, gender, and other atherogenic risk factors such as diabetes mellitus. We used calibrated Framingham equations as developed by Anderson et al. [13,14]. It has been known that the overall incidence/prevalence of CAD is very low in Japan, and it has been stressed that the Framingham prediction equation function should be recalibrated if it is to be used for populations with different characteristics. Therefore, we used the calibrated equation for the Japanese population data using the method of D'Agostino et al. [13,15,16]. Using data from the published literature, we determined the probability of falling in a particular disease category (1-, 2-, 3-vessel or left main trunk disease) by age group. We then assigned probabilities for myocardial ischemia or MI, and the probability that the initial myocardial ischemia would be asymptomatic [17–19].

As well as the 10 assumptions in the U.S. model [8,20–22], we made the following further assumptions: 1) diagnostic test performance does not vary between countries; 2) the benefit from interventions does not differ between two countries [23,24]. We also assumed that a diagnostic screening test was performed only once, at the first stage. Screening strategies included: 1) no screening; 2) if exercise electrocardiography was positive, it was followed by coronary angiography (CAG); 3) if exercise echocardiography was positive, it was followed by CAG; 4) if exercise SPECT was positive, it was followed by CAG. The performance of the screening tests was obtained from a meta-analysis of the diagnostic tests, which contains original studies from various countries including Japan [25].

2.2. Prognosis

Mortality rates were determined according to patient characteristics using all-cause mortality from the 19th Japanese life tables, together with: the standardized mortality ratio for patients with diabetes, the relative mortality ratio for the specific extent of disease from the CASS registry, and the mortality risk reduction with medication or interventions [8,19,20,26–28]. We used 1.5 as the standardized mortality ratio for patients with diabetes, derived from the U.S. population, because the mortality risk ratio for diabetes mellitus varies very little between seven different countries, including Japan [29,30].

The mortality risk ratio for the extent of CAD, and the mortality risk reduction by CABG, PTCA, aspirin, and simvastatin were derived from the published literature; these have been discussed in detail in our previous work [8,20,21,27,28,31–34]. We allowed the rates of nonfatal myocardial infarction and revascularization to vary depending on the initial treatment, and assumed that the risks of subsequent nonfatal myocardial infarction, PTCA or CABG depend on the extent of the CAD and the type of initial treatment [10].

We made the following assumptions concerning short-term intervention sequelae. CAG was associated with a 0.1% probability of death and a 0.06% probability of nonfatal myocardial infarction, regardless of the extent of the CAD [35]. PTCA was associated with mortality rates of 0.2% in 1-vessel disease and 0.9% in 2-vessel disease; the mortality rate of nonfatal myocardial infarction was 3.5% in one-vessel disease and 5.2% in 2-vessel disease [36]. CABG was associated with a mortality rate of 3.2%, and the probability of nonfatal myocardial infarction was 7.0% [37–39].

2.3. Costs and discounting

When doing a cost-effectiveness analysis it is important to relate medical charges/reimbursement to medical cost, because the two are different [40]. However, in Japan, for the cost of CAD care the cost/charge ratio is close to one [41]. Since the data are readily available, we therefore used medical charges based on the reimbursement schedule as a substitute for medical costs [42,43]. Data on other therapeutic procedures for CAD and the cost of conventional diabetes were derived from the reimbursement data of a public teaching hospital (Shimane Prefectural Central Hospital, Japan). We did not assign any CAD related costs for silent myocardial ischemia. The opportunity cost was estimated by including patient travel, waiting time, and treatment time associated with office visits, which were themselves estimated using survey data [1] and average hourly earnings of employed persons (average monthly earnings/monthly working hours) reported by the Ministry of Health, Labor and Welfare in 2003 [44]. All costs were adjusted to 2003 Japanese Yen using the medical care component of the Consumer Price Index [45], and converted to U.S. dollars using the OECD Purchasing Power Parity rate in 2003 (130.3 yen/dollar). All costs and years of life were discounted at 3% per year to reflect time preference [46].

2.4. Health utility

Health utility values, between 1 for perfect health and 0 for death, were used to calculate QALYs [47,48]. The utilities for symptomatic myocardial ischemia and myocardial infarction were derived from 10 patients with uncomplicated diabetes who were taking a diabetes educational course in a teaching

hospital (Tenri Hospital, Nara, Japan). We used the time trade-off method based on hypothetical clinical vignettes. Scenarios and example of time trade-off method we used to derive utilities were shown in Appendixes 1 and 2. A health utility of 1 was set for all other levels of being alive.

2.5. Sensitivity analysis

We performed one-way sensitivity analyses on all the variables within clinically plausible ranges. Table 1 shows the ranges used. Where applicable, 95% confidence intervals (CI) were used as the range of the variables in the sensitivity analysis; otherwise a ±30% range was used. There were two exceptions: we used the range of sensitivities and specificities of the diagnostic tests reported in the meta-analysis, since the 95% CIs reported in the meta-analysis were extremely narrow [25,49]; and we set the lower range of risk reduction for late myocardial infarction with PTCA at 0%, because in diabetic patients with asymptomatic 1- or 2-vessel disease there is not enough evidence to support a PTCA effect. The incidence of CAD in Japanese women increased by 36% over these 14 years, but has changed less in Japanese men. We therefore conducted a sensitivity analysis for women by increasing both the incidence and prevalence of CAD by 40%. We also conducted a two-way sensitivity analysis dealing with diagnostic performance [8].

In our base-case analysis, screening was performed only once, at the time of initial screening. To see if there was an optimal screening frequency, we performed sensitivity analyses for the most cost-effective strategy in the base case for screening occurring at intervals of every 3 years, every 5 years and every 10 years. Analyses were also conducted for cohorts of different ages and/or different pairs of additional atherogenic risk factors. We also evaluated the effect of varying the discount rate from 0% to 5% [46].

3. Results

3.1. Baseline analysis

Table 2 shows the quality-adjusted life expectancy, lifetime cost, and incremental cost-effectiveness ratio in asymptomatic 55- and 60-year-old diabetic men with hypertension and smoking. In 55-year-old men, compared to a no-screening strategy, the incremental cost-effectiveness ratio of exercise electrocardiography was $61,300/QALY, and that of exercise echocardiography was $61,100/QALY. In 60-year-old men, compared to not screening, the incremental cost-effectiveness ratio of exercise electrocardiography was $31,400/QALY. Compared with electrocardiography, exercise echocardiography cost $31,500 per QALY saved. The exercise SPECT strategy had a higher cost and small incremental benefit, and

Table 2
Quality-adjusted life years, cost, and cost-effectiveness ratios for asymptomatic men and women with hypertension and smoking*

Screening strategy	Expected value		Incremental value		Incremental cost-effectiveness ratio*
	Cost	QALYs	Cost	QALYs	
	$	Y	$	y	$/y
55-year-old men					
No screening	150,545	13.2311			
Exercise electrocardiography	154,086	13.2889	(3541)	(0.0578)	(61,300)
Exercise echocardiography	154,989	13.3038	4444	0.0727	61,100
Exercise SPECT	155,770	13.3042	781	0.0004	195,000
60-year-old men					
No screening	135,332	11.2402			
Exercise electrocardiography	138,986	11.3566	3654	0.1164	31,400
Exercise echocardiography	139,917	11.3862	931	0.0296	31,500
Exercise SPECT	140,699	11.3886	782	0.0024	326,000
65-year-old women					
No screening	105,275	11.6932			
Exercise electrocardiography	108,519	11.7628	(3244)	(0.0696)	(46,600)
Exercise echocardiography	109,348	11.7808	4073	0.0876	46,500
Exercise SPECT	110,119	11.7817	771	0.0009	857,000
70-year-old women					
No screening	87,166	9.0992			
Exercise electrocardiography	91,207	9.2498	4041	0.1506	26,800
Exercise echocardiography	92,236	9.2878	1029	0.0380	27,100
Exercise SPECT	93,025	9.2914	789	0.0036	219,000

*Incremental cost-effectiveness ratios for each strategy are calculated compared with the next strategy other than dominated strategy shown in the table, and are rounded to the nearest $100. (Note: Cost-effectiveness ratios calculated directly by quality-adjusted life expectancies and costs from the table may differ due to rounding.)

consequently cost $326,000 per QALY saved, compared with echocardiography.

3.2. Patients with different characteristics

The incremental cost-effectiveness varied with age, gender, and the pairing of additional atherogenic risk factors. The cost-effectiveness of exercise echocardiography relative to a no-screening strategy was sensitive to the age of patients at the time of screening. It fell from $61,100/QALY to $17,600/QALY as the age at initial screening rose from 55 to 70 years. In regard to gender, the incremental cost-effectiveness of echocardiography was $46,600/QALY in 65-year-old women with HTN and smoking, falling to $26,800 in 70-year-old women (Table 2). These results were stable even when the incidence and the prevalence both increased by 40%. Fig. 1 shows incremental cost-effectiveness ratios for exercise echocardiography and exercise electrocardiography compared with the next dominant strategy for 60-year-old men and 65-year-old women, with three possible pairs of additional atherogenic risk factors. Our model was not sensitive to the differences in the pairs of risk factors; depending on the additional atherogenic risk factors, the incremental cost-effectiveness ratio of exercise echocardiography compared with electrocardiography in 60-year-old men ranged from $31,200/QALY (high LDL/hypertension) to $31,400/QALY (hypertension/smoking).

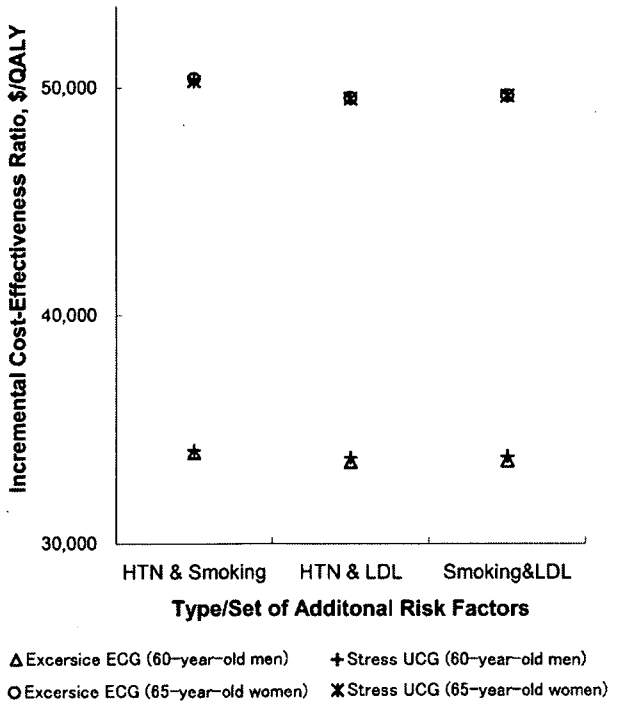

Fig. 1. Cost-effectiveness ratio for alternative patient cohorts. Triangles (Δ) represent exercise electrocardiography compared with no screening in 60-year-old men; plus signs (+) represent exercise echocardiography compared with exercise electrocardiography in 60-year-old men; circles (O) represent exercise electrocardiography compared with no screening in 65-year-old women; asterisks (*) represent exercise echocardiography compared with no testing in 65-year-old women. SPECT indicates single-photon emission-tomography; QALY, quality-adjusted life expectancy; LDL, low-density lipoprotein; HTN, hypertension; ECG, electrocardiography; UCG, echocardiography.

Table 3
Incremental cost-effectiveness of exercise echocardiography vs. no screening at different frequencies of screening*

60-year-old men	Expected value		Incremental value		Incremental cost-effectiveness ratio*
	Cost	QALYs	Cost	QALYs	
	$	y	$	y	$/y
Screening strategies					
No screening	135,332	11.2402			
Only at initial screening	139,917	11.3862	4585	0.144	31,800
Every 10 years	145,155	11.3912	5238	0.005	1,047,600
Every 5 years	152,983	11.3986	7828	0.0074	1,057,800
Every 3 years	162,792	11.4040	9814	0.0054	1,817,400

*Incremental cost-effectiveness ratios for each strategy are calculated compared with the next strategy other than dominated strategy shown in the table, and are rounded to the nearest $100. (Note: Cost-effectiveness ratios calculated directly by quality-adjusted life years and costs from the table may differ due to rounding.)

3.3. Sensitivity analysis

The incremental cost-effective ratios of exercise electrocardiography or echocardiography, compared in 60-year-old men, did not exceed $50,000/QALY in the sensitivity analyses to the following variables: mortality risk reduction by CABG or PTCA; risk reduction of myocardial infarction by CABG or PTCA; the cost of CABG; the cost of screening tests; the proportion of patients with silent myocardial ischemia; diagnostic test performance; or the health utility of CAD.

For 60-year-old men with diabetes and two atherogenic risk factors, the incremental cost-effectiveness ratio for echocardiography increased as the interval of screening became shorter (Table 3). The incremental cost-effectiveness ratio increased from $31,800 in the base case to $1,060,000/QALY if screening was done every 5 years, and to $1,820,000/QALY if screening was done every 3 years.

4. Discussion

We have shown that the U.S. guidelines for screening for CAD in high risk asymptomatic diabetic patients may be applicable to the Japanese population in cost-effectiveness terms. We also found that implementation of any screening should take into account patient age, gender, and the frequency of the tests. In our previous study, screening for CAD with exercise echocardiography was cost-effective in men over 60 years of age in the U.S., with an incremental cost-effectiveness ratio of $40,800/QALY [8]; this fell to $31,500 in Japanese men of the same age with the same pair of risk factors. Our results might be affected by the incidence/prevalence of CAD in the Japanese population. There is known to be international variation in the incidence/

prevalence of CAD. The overall prevalence of CAD in Japan has been estimated at 6.3 per 1000 persons, which is much lower than the value of 46 per 1000 in the U.S. [2,26]. The main reason for the difference could be the lower prevalence of obesity and dyslipidemia in the Japanese population. However, the incidence/prevalence of CAD in Japanese diabetic patients is only slightly below that in American diabetic patients with similar risk factors, which could explain the higher cost per QALY saved in Japan. Cost may be a further influence; the lower average cost of diagnostic tests and other medical costs in Japan may explain why screening for CAD costs less per QALY in Japan than in the U.S.

Not surprisingly, the incremental cost-effectiveness ratio of screening for CAD in women was much higher than in men of the same age. The incremental cost-effectiveness of screening with echocardiography in 60-year-old women was $62,600/QALY, compared to $31,500 in men of the same age. The main reason for this cost differential is that, in Japan, the incidence and prevalence of CAD is much lower in women than in men having the same pair of risk factors. Also, the life expectancy of Japanese women is the world's greatest at 85.3 years, so that their baseline mortality rate is also lower than for men [50]. The mortality risk for invasive diagnostic tests and therapeutic procedures was therefore likely to exceed the benefit derived from screening. Furthermore, a sensitivity analysis found that these results were not sensitive to changes in incidence and prevalence liable to occur in the next 10 to 15 years. Consequently, we recommend screening for CAD in high risk diabetic women over the age of 65–70 years.

An incremental cost-effectiveness ratio higher than $50,000/QALY generally exceeds the currently accepted level for health technologies. However, the actual criteria for acceptable cost-effectiveness ratio appear to depend strongly on social circumstances, so that the $50,000/QALY threshold should not be necessarily applied to our results. The incremental cost-effectiveness ratio for gastric cancer screening in men was $5800 per life expectancy (LE), and in women it was $14,800/LE [51]; it was found to be $26,800/LE for breast cancer screening by physical examination and mammography [52]. These screening programs have already been accepted and implemented across Japan for many years, so that the incremental cost-effectiveness of CAD screening seems to be within the acceptable range.

Application of evidence-based guidelines from one country within another country demands careful consideration. A tenet of evidenced-based medicine (EBM) is to do the right things right for the right people at the right time [53]. However, finding out what is right is influenced not only by scientific fact, but also by political, economic, and sociocultural factors. These may influence when care can or should be delivered to a particular population, and what type of care. We may draw two conclusions from the present results. First, when dealing with screening strategies for a high risk population, country-specific factors appear to play less of a role. However, U.S. guidelines on the use of aspirin for the primary prevention of CAD were of limited use in Japan where there was a high incidence of hemorrhagic stroke and gastrointestinal bleeding [54]. Since screening tests generally do less harm than therapeutic agents, fewer factors should limit the applicability of screening guidelines. Also, after controlling for the baseline risks of the target population, cost became an influential factor on the cost-effectiveness of applying the guidelines, and should therefore be taken into account when considering the adoption of foreign guidelines.

When determining optimal screening frequencies, it is necessary to consider both the clinical benefits that patients may gain and the added cost [55]. In general, increasing the frequency of screening results in an unfavorable incremental cost-effective ratio. For example, when screening individuals with neither hypertension nor diabetes for proteinuria, the incremental cost-effectiveness ratio increased from $80,700/QALY to $120,727/QALY as the screening frequency increased from every 10 years to every 5 years [56]. Similar patterns are visible for diabetic retinopathy screening or abdominal aortic aneurysm screening [55,57]. In our model, individuals who had not been detected at the initial screening and were re-screened had a lower risk of CAD than those identified in the initial cohort. The cost-benefit ratio of screening might thereby have increased despite an expected higher incidence rate with increased age. Guidelines have not explicitly mentioned the frequency of screening, but groups making recommendations on the frequency of screening should decide whether the clinical benefit of screening CAD outweighs the added cost of more frequent screening.

Conducting clinical studies in differing countries is a costly and time-consuming way to evaluate how health resources should be invested so as to maximize clinical benefits and minimize costs. A mathematical simulation model can be a useful tool for evaluating the applicability of diagnostic/therapeutic strategies to different economic/cultural situations, by extending the knowledge already gained from local studies. An example was a study conducted to determine the optimal screening strategy for cervical cancer in South Africa, which demonstrated how mathematical methods could connect research and health policy [7]. Cost-effectiveness analyses are seldom used for these purposes, but are a promising way to save on investment.

Cost-effectiveness analyses assist both in assessing the economic impact of a significant intervention, such as screening for CAD in patients with diabetes, and in identifying areas of uncertainty so as to improve decision making [62]. The present analysis has revealed uncertainty surrounding probability estimates. The main area of uncertainty lay in the data dealing with the incidence and prevalence of CAD; these data are still sparse in Japan, especially data stratified by gender, age, or other atherogenic risk factors. This lack of data could have been a limiting factor in our study. Furthermore, the size of the cohort that we used to recalibrate the Framingham prediction equation was very small, which could have biased our estimates. We also had to calculate the

prevalence of CAD based on the relative risk estimated from a cohort of employees [12]. Since the subjects of this study were relatively young, our estimate of the prevalence of CAD in higher age groups might have been biased. More detailed data should be made publicly available in order to improve the quality of cost-effectiveness analysis in Japan.

A limitation of this study is that we derived some estimates from the studies published more than ten years ago, but we believe that our analyses nevertheless remain valid for the following reasons. First, over these ten years, for single or two-vessel disease, a metal stent has come into use, and the revascularization rate at one year of 26.4% in a recent study is lower than the estimate used for the current analysis (36%); this value of 26.4% is nevertheless within the range of sensitivity analysis. Second, with regard to drug eluting stent, this has recently been approved in Japan by the Ministry of Health and Welfare; this modality is not yet used widely in Japan, so that there is insufficient cost of data for deployment in the present analysis. Third, for multivessel or left main trunk disease, several old studies comparing CABG vs. PTCA in subgroups of diabetic patients demonstrated a survival advantage and fewer repeat revascularization procedures with an initial surgical strategy. Advances in medical therapy (e.g., drug eluting stent) and surgical techniques need to be included in the analysis; however, no reliable clinical trial has been done. Two current RCTs sponsored by the National Heart, Lung, and Blood institutes of Health compare PCI and CABG surgery in diabetic patients (BARI 2 Diabetes, FREEDOM), and their results could cause us to modify the present model if the effect size of these interventions proves to be out of range in our sensitivity analyses.

Finally, few patients with diabetes (~6%) were included in the Framingham cohort [58], so that this equation is not specifically designed for patients with diabetes; it also does not take into account the duration of diabetes or glycemic control. The UKPDS risk engine is a diabetes-specific CAD risk assessment tool, with the duration of diabetes and glycemic control incorporated [59]. We nevertheless used the Framingham risk equation to estimate the CAD risk, for the following reasons. First, it was validated and calibrated for the Japanese population in previous studies, unlike the UKPDS risk engine. Second, it has been reported that these two methods are comparable in identifying those at high risk of CAD [60]. Third, the mean duration of diabetes in the Framingham cohort was 7.8 years [61], and our results could be applicable to those with a similar duration of diabetes. It would still be useful to validate the UKPDS risk engine for the Japanese population, and a group with a clear potential benefit, e.g. duration of diabetes or glycemic control, could be specified.

In conclusion, from a societal perspective, U.S. guidelines on screening for CAD in high risk diabetic patients are applicable to the Japanese population. However, the actual population that could benefit from these strategies depends on the patient characteristics and the diagnostic performance of the screening tests. It is therefore still necessary to select carefully the target population for screening by considering gender, age, type of atherogenic risk factors, and the frequency of tests.

Acknowledgements

Hayashino Y. received fellowship grants from St. Luke's Life Science Institute, Tokyo.

Appendix A. Clinical scenarios used to derive utilities in the time trade-off method

A.1. Symptomatic myocardial ischemia

From about 6 months ago, you began to feel chest oppression every time you did exercise. The chest pain was relieved when you stopped doing exercise. You consulted your primary care physician, and he referred you to a cardiologist because the symptom was consistent with ischemic heart disease. The cardiologist recommended your admittance, and you underwent coronary angiography. This is an invasive diagnostic procedure, in which a very small gauge tube is inserted from the femoral artery to your heart. It can test whether or not your coronary arteries, which supply oxygen to your heart muscles and are critical to your heart function, are occlusive. You doctor explained to you that one of your three coronary arteries which supplies your heart muscle was narrowed, and explaining your chest oppression when you did exercise. You therefore decided to undertake therapy to expand the occluded vessels by inserting a small balloon into the vessel. After you were discharged, you no longer felt chest oppression even you do exercise. But after around 6 months after discharge, you began to feel the same symptom again when you did exercise. Your cardiologist told you that your vessel may be narrowed again, and prescribed medication to expand your vessels and to stop your blood from clotting. You do not feel any constraint in everyday life, but you sometimes feel chest oppression when you do exercise or if you feel tense, or on very cold days. It is not perfectly bothersome, but you feel uneasy because you do not know when it will occur. Your cardiologist advised you that you may be admitted and undergo cardiac catheterization if you are bothered by your chest pain, but you feel that your chest pain is not so bad that have you make that choice.

A.2. History of myocardial infarction

One month ago, while reading a book at home, you suffered the sudden onset of chest pain, which was the strongest you had experienced. You next felt dyspnea soon after the onset of chest pain, and you had your family call emergency medical service to visit ER. You underwent on initial evaluation including blood test, electrocardiography, and chest X-ray testing, which revealed that you had acute myocardial infarction and congestive heart failure. You underwent coronary angiography. This is an invasive

diagnostic procedure, in which a very small gauge tube is inserted from the femoral artery to your heart. It can test whether or not your coronary arteries, which supply oxygen to your heart muscles and are critical to your heart function, are occlusive. Coronary angiography revealed that one of the three coronary arteries was completely occlusive, and that as a result one-third of your heart muscle was going to die and the function of the heart as a pump would deteriorate, which was why you felt dyspnea. You therefore undertook therapy to expand the occluded vessels by inserting small balloon into the vessel. Your chest pain was relieved by the therapeutic procedure, but your heart function remained reduced after suffering from myocardial infarction. After several days, you were discharged from the hospital. You periodically see a cardiologist, and are taking drugs including aspirin and nitrate which prevent your coronary vessels from occluding. At usual daily life, you do not feel chest pain or dyspnea, but you feel slight dyspnea when you walk longer distances.

Appendix B. Example of the time trade-off method used for the current study

Assume that your natural life expectancy is 10 years with chronic stable angina. Now imagine that someone who has supernatural power offers you a different magic pill. If you take this pill, you will be relieved of your chronic stable angina forever, but you live only 5 years, instead of 10. Would you take this pill? If your answer is no, would you take the pill if it would relieve your chronic stable angina and you would live 8 more years? If your answer is yes, would you take the pill if it would relieve your heart disease and you would live 6 years?

References

[1] Statistics and information Department, Ministers Secretariat, Ministry of Health and Welfare, patient behavioral survey 1999. Tokyo, Japan: Ministry of Health and Welfare; 2001. [in Japanese].

[2] American Heart Association. Heart Disease and Stroke Statistics-2004 Update.

[3] Ministry of Health and Welfare. 2002 survey on diabetes. Tokyo, Japan: Ministry of Health and Welfare; 2003. [in Japanese].

[4] Jacoby RM, Nesto RW. Acute myocardial infarction in the diabetic patient: pathophysiology, clinical course and prognosis. J Am Coll Cardiol 1992;20(3):736–44.

[5] Fava S, Azzopardi J, Muscat HA, Fenech FF. Factors that influence outcome in diabetic subjects with myocardial infarction. Diabetes Care 1993;16(12):1615–8.

[6] Consensus development conference on the diagnosis of coronary heart disease in people with diabetes: 10–11 February 1998, Miami, Florida. American Diabetes Association. Diabetes Care 1998;21(9):1551–9.

[7] Goldie SJ. Health economics and cervical cancer prevention: a global perspective. Virus Res 2002;89(2):301–9.

[8] Hayashino Y, Nagata-Kobayashi S, Morimoto T, Maeda K, Shimbo T, Fukui T. Cost effectiveness of screening for coronary artery disease in asymptomatic patients with type 2 diabetes and additional atherogenic risk factors. J Gen Intern Med 2004;19:1181–91.

[9] Cohn PF. Silent myocardial ischemia. Ann Intern Med 1988;109 (4):312–7.

[10] Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. Med Decis Making 1993;13(4):322–38.

[11] Weinstein MC, Coxson PG, Williams LW, Pass TM, Stason WB, Goldman L. Forecasting coronary heart disease incidence, mortality, and cost: the Coronary Heart Disease Policy Model. Am J Public Health 1987;77(11):1417–26.

[12] Nakamura T, Tsubono Y, Kameda-Takemura K, et al. Magnitude of sustained multiple risk factors for ischemic heart disease in Japanese employees: a case-control study. Jpn Circ J 2001;65(1):11–7.

[13] D'Agostino Sr RB, Grundy S, Sullivan LM, Wilson P. Validation of the Framingham coronary heart disease prediction scores: results of a multiple ethnic groups investigation. JAMA 2001;286(2):180–7.

[14] Anderson KM, Odell PM, Wilson PW, Kannel WB. Cardiovascular disease risk profiles. Am Heart J 1991;121(1 Pt 2):293–8.

[15] Nomura H, Koizumi J, Hayashi T, Matsukura T, Mabuchi H. Annual meeting for the Japanese society of general medicine, Kagawa, Japan; 2002. p. 43. [in Japanese].

[16] Koizumi J, Shimizu M, Miyamoto S, Origasa H, Mabuchi H. Effect of pravastatin-induced LDL-cholesterol reduction on coronary heart disease and cerebrovascular disease in Japanese: Hokuriku lipid coronary heart disease study-pravastatin atherosclerosis trial (Holicos-PAT). J Atheroscler Thromb 2002;9(5):251–9.

[17] Hunink MG, Goldman L, Tosteson AN, et al. The recent decline in mortality from coronary heart disease, 1980–1990. The effect of secular trends in risk factors and treatment. JAMA 1997;277 (7):535–42.

[18] Airaksinen KE. Silent coronary artery disease in diabetes—a feature of autonomic neuropathy or accelerated atherosclerosis? Diabetologia 2001;44(2):259–66.

[19] Chaitman BR, Bourassa MG, Davis K, et al. Angiographic prevalence of high-risk coronary artery disease in patient subsets (CASS). Circulation 1981;64(2):360–7.

[20] Shepherd J, Cobbe SM, Ford I, et al. Prevention of coronary heart disease with pravastatin in men with hypercholesterolemia. West of Scotland Coronary Prevention Study Group. N Engl J Med 1995;333(20):1301–7.

[21] Collins R, Armitage J, Parish S, Sleigh P, Peto R. MRC/BHF Heart Protection Study of cholesterol-lowering with simvastatin in 5963 people with diabetes: a randomised placebo-controlled trial. Lancet 2003;361(9374):2005–16.

[22] Haffner SM. Dyslipidemia management in adults with diabetes. Diabetes Care 2004;27(Suppl 1):S68–71.

[23] Shihara M, Tsutsui H, Tsuchihashi M, Tada H, Kono S, Takeshita A. In-hospital and one-year outcomes for patients undergoing percutaneous coronary intervention for acute myocardial infarction. Am J Cardiol 2002;90(9):932–6.

[24] Mastoor M, Iqbal U, Pinnow E, Lindsay Jr J. Ethnicity does not affect outcomes of coronary angioplasty. Clin Cardiol 2000;23(5):379–82.

[25] Gianrossi R, Detrano R, Mulvihill D, et al. Exercise-induced ST depression in the diagnosis of coronary artery disease. A meta-analysis. Circulation 1989;80(1):87–98.

[26] Statistics and information Department, Ministers Secretariat, Ministry of Health and Welfare, patient survey 1999. Tokyo, Japan: Ministry of Health and Welfare; 2001. [in Japanese].

[27] Yusuf S, Zucker D, Peduzzi P, et al. Effect of coronary artery bypass graft surgery on survival: overview of 10-year results from randomised trials by the Coronary Artery Bypass Graft Surgery Trialists Collaboration. Lancet 1994;344(8922):563–70.

[28] Collaborative overview of randomised trials of antiplatelet therapy-I: prevention of death, myocardial infarction, and stroke by prolonged antiplatelet therapy in various categories of patients. Antiplatelet Trialists' Collaboration. BMJ 1994;308(6921):81–106.

[29] Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. Circulation 1998;97(18):1837–47.

[30] Menotti A, Blackburn H, Seccareccia F, et al. The relation of chronic diseases to all-cause mortality risk—the seven countries study. Ann Med 1997;29(2):135–41.

[31] Mark DB, Nelson CL, Califf RM, et al. Continuing evolution of therapy for coronary artery disease. Initial results from the era of coronary angioplasty. Circulation 1994;89(5):2015–25.
[32] Pocock SJ, Henderson RA, Rickards AF, et al. Meta-analysis of randomised trials comparing coronary angioplasty with bypass surgery. Lancet 1995;346(8984):1184–9.
[33] Sanmuganathan PS, Ghahramani P, Jackson PR, Wallis EJ, Ramsay LE. Aspirin for primary prevention of coronary heart disease: safety and absolute benefit related to coronary risk derived from meta-analysis of randomised trials. Heart 2001;85(3):265–71.
[34] Gaspoz JM, Coxson PG, Goldman PA. Cost effectiveness of aspirin, clopidogrel, or both for secondary prevention of coronary heart disease. N Engl J Med 2002;346(23):1800–6.
[35] Johnson LW, Lozner EC, Johnson S. Coronary arteriography 1984–1987: a report of the Registry of the Society for Cardiac Angiography and Interventions. I. Results and complications. Cathet Cardiovasc Diagn 1989;17(1):5–10.
[36] Holmes Jr DR, Holubkov R, Vlietstra RE. Comparison of complications during percutaneous transluminal coronary angioplasty from 1977 to 1981 and from 1985 to 1986: the National Heart, Lung, and Blood Institute Percutaneous Transluminal Coronary Angioplasty Registry. J Am Coll Cardiol 1988;12(5):1149–55.
[37] Yusuf S, Zucker D, T.C.C. Ten-year results of the randomized controlled trials of coronary artery bypass graft surgery: tabular data complied by the collaborative effort of the original trial investigators. Part 1 of 2. Online J Curr Clin Trials 1994;14 [document no.144].
[38] Yusuf S, Zucker D, T.C.C. Ten-year results of the randomized controlled trials of coronary artery bypass graft surgery: tabular data complied by the collaborative effort of the original trial investigators. Part 2 of 2. Online J Curr Clin Trials 1994;14 [document no.145].
[39] Hannan EL, Kilburn Jr H, O'Donnell JF, Lukacik G, Shields EP. Adult open heart surgery in New York State. An analysis of risk factors and hospital mortality rates. JAMA 1990;264(21):2768–74.
[40] Finkler SA. The distinction between cost and charges. Ann Intern Med 1982;96(1):102–9.
[41] Quality-of-life Policy Bureau. The report from working party for medical costs. announcement in 07 March, 2001 Available from URL: http://www5.cao.go.jp/seikatsu/2001/0307iryou-houkoku/menu.html [in Japanese].
[42] Division of Medicine, Department of Insurance, Japanese Ministry of Health and Welfare. Interpretation of the social insurance reimbursement table. Tokyo: Research Institute of Social Insurance; 2000. (in Japanese).
[43] Nagata-Kobayashi S, Shimbo T, Fukui T. Cost-effectiveness analysis of screening for osteoporosis in postmenopausal Japanese women. J Bone Miner Metab 2002;20(6):350–7.
[44] Ministry of Health, Labour and Welfare. White Paper on the Labour Economy 2003 Japan Institute of Labour, Tokyo [in Japanese].
[45] Statistics Bureau (2003) Japan Statistical Yearbook, 2004. Ministry of Public Management, Home Affairs, Posts, and Telecommunications, Tokyo. [in Japanese].
[46] Weinstein MC, Siegel JE, Gold MR, Kamlet MS, Russell LB. Recommendations of the panel on cost-effectiveness in health and medicine. JAMA 1996;276(15):1253–8.
[47] Torrance GW. Preferences for health states: a review of measurement methods. Mead Johnson Symp Perinat Dev Med 1982;20:37–45.
[48] Torrance GW, Thomas WG, Sackett DL. A utility maximization model for evaluation of health care programs. Health Serv Res 1972;7:118–33.
[49] Fleischmann KE, Hunink MG, Kuntz KM, Douglas PS. Exercise echocardiography or exercise SPECT imaging? A meta-analysis of diagnostic test performance. JAMA 1998;280(10):913–20.
[50] Statistics and Information Department, Ministers Secretariat, Ministry of Health and Welfare. Abridged life table 2003. Tokyo, Japan: Ministry of Health and Welfare; 2003. [in Japanese].
[51] Tsuji I, Fukao A, Sugawara N, Shoji T, Kuwajima I, Hisamichi S. Cost-effectiveness analysis of screening for gastric cancer in Japan. Tohoku J Exp Med 1991;164(4):279–84.
[52] Okubo I, Glick H, Frumkin H, Eisenberg JM. Cost-effectiveness analysis of mass screening for breast cancer in Japan. Cancer 1991;67 (8):2021–9.
[53] Reilly BM. The essence of EBM. BMJ 2004;329(7473):991–2.
[54] Morimoto T, Fukui T, Lee TH, Matsui K. Application of U.S. guidelines in other countries: aspirin for the primary prevention of cardiovascular events in Japan. Am J Med 2004;117(7):459–68.
[55] Vijan S, Hofer TP, Hayward RA. Cost-utility analysis of screening intervals for diabetic retinopathy in patients with type 2 diabetes mellitus. JAMA 2000;283(7):889–96.
[56] Boulware LE, Jaar BG, Tarver-Carr ME, Brancati FL, Powe NR. Screening for proteinuria in US adults: a cost-effectiveness analysis. JAMA 2003;290(23):3101–14.
[57] Frame PS, Fryback DG, Patterson C. Screening for abdominal aortic aneurysm in men ages 60 to 80 years. A cost-effectiveness analysis. Ann Intern Med 1993;119(5):411–6.
[58] Yeo WW, Yeo KR. Predicting CHD risk in patients with diabetes mellitus. Diabet Med 2001;18(5):341–4.
[59] Stevens RJ, Kothari V, Adler AI, Stratton IM. The UKPDS risk engine: a model for the risk of coronary heart disease in Type II diabetes (UKPDS 56). Clin Sci (Lond) 2001;101(6):671–9.
[60] Song SH, Brown PM. Coronary heart disease risk assessment in diabetes mellitus: comparison of UKPDS risk engine with Framingham risk assessment function and its clinical implications. Diabet Med 2004;21(3):238–45.
[61] Fox CS, Sullivan L, D'Agostino Sr RB, Wilson PW. The significant effect of diabetes duration on coronary heart disease mortality: the Framingham Heart Study. Diabetes Care 2004;27(3):704–8.
[62] Hunink M, Glasziou P, Siegel J, et al. Decision making in health and medicine: integrating evidence and values. Cambridge: Cambridge University Press; 2001.
[63] Alexander CM, Landsman PB, Teutsch SM. Diabetes mellitus, impaired fasting glucose, atherosclerotic risk factors, and prevalence of coronary heart disease. Am J Cardiol 2000;86(9):897–902.

International Journal of Urology (2007) 14, 805–810 doi: 10.1111/j.1442-2042.2007.01824.x

Original Article: Clinical Investigation

Impact of improvement in specificity of primary screening test on total cost of prostate cancer mass screening

Takashi Kobayashi,[1,2] Rei Goto,[3] Tsuguya Fukui[4] and Osamu Ogawa[1]

[1]Department of Urology, Kyoto University Graduate School of Medicine, Kyoto, [2]Japan Society for the Promotion of Science (JSPS), Tokyo, [3]Faculty of Economics, Konan University, Kobe, and [4]St. Luke's International Hospital, Tokyo, Japan

Objectives: Improvement in the specificity of primary screening tests, without significant cost elevation of the assay, provides cost savings in prostate cancer screening programs by reducing unnecessary secondary screening procedures. The purpose of this study was to evaluate the economic impact of improvement in the specificity of primary screening tests and to estimate the socially acceptable cost elevation for improved specificity.

Methods: A decision-analytic model was designed to evaluate the total costs of prostate cancer mass screening according to the changes in the specificity and the cost of the primary screening test. All assumed factors were tested by three-way sensitivity analyses incorporating cost and specificity.

Results: The base case analysis showed that a 1% improvement in the specificity of the primary screening test provides a $1.19 cost reduction per participant. Sensitivity analyses showed that an acceptable cost elevation for a 1% improvement in the specificity ranged from 0.68 to 2.90 $/% with respect to changes in several factors in the screening program.

Conclusions: The specificity and cost of the primary screening test has a significant economical impact on prostate cancer mass screening. For each screening program, it should be taken into consideration whether the cost of the new test deserves the specificity.

Key words: decision-analytic model, prostate-specific antigen, sensitivity analysis.

Introduction

Worldwide, prostate cancer is the third most common cancer and the cause of 6% of all cancer deaths in men.[1] It is the most frequently diagnosed cancer and the second leading cause of cancer deaths in men in Western countries.[2,3] Since treatment outcome of prostate cancer depends highly on the disease stage, early detection is an important strategy in reducing prostate cancer mortality. On the strength of prostate-specific antigen (PSA) testing, a convenient and less invasive method for screening patients who need further diagnostic examination, prostate cancer is considered one of the most suitable cancers for mass screening. Although some investigators strike a note of warning against the current popularization of prostate cancer mass screening without definitive evidence of reducing prostate cancer mortality,[4] a recent report demonstrated that early detection of prostate cancer by PSA-based serial screening effectively decreases prostate cancer-specific mortality.[5]

Although PSA is one of the most useful tumor markers, the specificity in primary screening of prostate cancer is not sufficiently high. One of the most serious concerns about PSA-based screening is that widespread use of a less specific primary screening results in an increased number of false positives and consequent unnecessary biopsies in men without cancer. It has been an urgent problem to improve the specificity of the primary screening by incorporating additional information including age, molecular forms, prostate volume and digital examination findings, or developing a new assay replaceable for conventional PSA. In the recent studies the relative merits of primary screening measures for prostate cancer detection were determined by their specificity at sensitivity levels of 80–95%.[6]

The less specific nature of PSA resulting in increased false positive results and unnecessary secondary screening procedures can also cause an increased cost burden as well as impaired quality of life for participants. On the contrary, the cost of secondary screening procedures (biopsy) would be saved if a new primary screening measure with higher specificity were introduced. However, the impact of the increase in cost of primary screening should be also considered if the new primary screening measure is more expensive than the conventional one. The cost elevation of the primary screening should be kept under an acceptable degree that does not overbalance the reduction in cost of secondary screening provided by the improved specificity. Eventually, a new primary screening measure could be introduced to the screening program if the reduced cost of secondary screening yielded from the improved specificity exceeds the increase in primary screening cost, whereas less specific or more expensive screening measures cannot be used.

Theoretically, there must be a threshold of acceptable cost elevation for a certain degree of improvement in the specificity and this can be determined by analysing the impact of the specificity improvement on the total screening cost. It would be helpful for both promoters of screening programs and developers of primary screening measures. However, there has been little evidence on the concept of acceptable cost elevation for primary screening measures for prostate cancer detection. Here, using a simple decision-analytic model for prostate cancer mass screening, we simulated the cost function for evaluating the impact of the improvement in the specificity with the

Correspondence: Osamu Ogawa MD, Department of Urology, Kyoto University Graduate School of Medicine, Shogoinkawahara-cho 54, Sakyo-ku, Kyoto, 606-8507, Japan. Email: ogawao@kuhp.kyoto-u.ac.jp
Contributions: TK 45%, RG 30%, TF 10%, OO 15%.

Received 29 January 2007; accepted 11 April 2007.

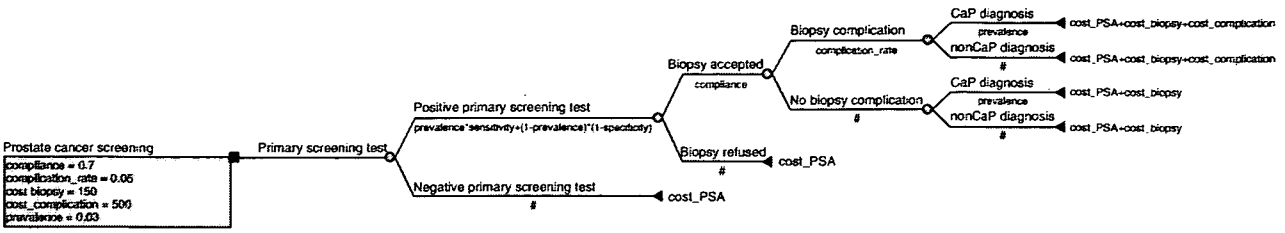

Fig. 1 Decision-analytic model of prostate cancer screening program. CaP: prostate cancer; cost_PSA: cost of prostate-specific antigen; cost_biopsy: cost of prostate biopsy; cost_complication: cost of the management of complication from prostate biopsy.

Table 1 Large scale prostate cancer screening programs

Program	PLCO[7]	PLCO[7]	Prostate cancer awareness week[8]	Prostate cancer awareness week[8]	PSA-2 study[9]	ERSPC Rotterdam section[10]	Tyrol study[11]	Gunma Study[12]
Period	1993–2001	1993–2001	1993–1994	1993–1994	1991–1996	1993–2001	1993–1994	1994–1999
Age	55–74	55–74	50–93	50–93	50–95	55–74	45–79	60–
Country	USA	USA	USA	USA	USA	the Netherlands	Austria	Japan
Number of participants	34 285	34 050	31 953	31 953	19 421	11 411	21 078	15 407
Primary screening test	PSA	PSA, DRE	PSA	PSA, DRE	PSA, DRE	PSA	PSA	PSA, DRE, TRUS
% Positive screening test	7.9	14.1	9.7	19.2	16.7	11.0	8.0	8.2
Biopsy compliance (%)	40.9	31.5	24.9	21.3	77.6	93.0	48.0	59.5
Positive predictive value	0.44	0.368	0.316	0.246	0.248	0.280	0.253	0.214
Overall cancer detection rate	1.4	1.4	3.6	4.7	3.2	3.1	0.9	1.8

% Positive screening test: Proportion of men with a positive result to a primary screening test of all participants. Biopsy compliance: Proportion of men undergoing biopsy to men with a positive primary screening test. Positive predictive value: Proportion of men with a positive biopsy result to men undergoing biopsy. Overall cancer detection rate: Proportion of men with a positive biopsy result to all participants. DRE, digital rectal examination; ERSPC, European Randomized Study of Screening for Prostate Cancer; PLCO, Prostate, Lung, Colorectal and Ovarian; PSA, prostate-specific antigen; TRUS, transrectal ultrasonography.

adjusted sensitivity of the primary screening on the total screening cost.

Methods

A decision-analytic model of a prostate cancer screening program was created (Fig. 1). Outcomes for the decision-analytic model were expressed in the expected total screening cost for each participant. In the model, participants undergo primary screening first, and secondary screening (biopsy) is recommended if the participants have a positive result of primary screening. The number of participants having secondary screening procedures after a positive primary screening result depends on disease prevalence, the sensitivity and specificity of the primary screening test, compliance of participants and other factors.

Probabilities were assumed based on the results of practical prostate cancer screening programs worldwide (Table 1) and the published reports (Table 2). In the base case, the proportion of men undergoing biopsy to men with positive primary screening tests (biopsy compliance) was assumed to be 70% according to reports on prostate cancer

Table 2 Assumption of clinical variables used in the base case and sensitivity analyses of the decision-analytic model

	Value	Reference
Biopsy compliance (%)	70	Table 1
% Complications of biopsy	5	13–15
Cost of biopsy ($)	150	16
Cost of management for complication of biopsy ($)	500	17
Sensitivity of primary screening tests (%)	95.0	6
Specificity of primary screening tests (%)	10.0	6
Prostate cancer prevalence	0.03	18,19, Table 1