

## 本日の検討の進め方

### 1. 本日の議論について

- 平成24年6月27日の本部会（第2回）において、当面の検討事項とされた「具体的な評価の運用手法（下図2.（2）」のうち、「① 2」効果指標の取り扱い」について、引き続き、検討を行うこととする。

当面の検討事項（太字下線部は、本日の検討事項）

（平成24年6月27日 費-3 より一部改変して抜粋）

#### （1）制度の基本的考え方

- ① 対象技術の原則
- ② 結果活用の原則

#### （2）具体的な評価の運用手法

##### ① 評価手法

- 1) 費用の範囲や取り扱い
- 2) **効果指標の取り扱い**
- 3) 比較対照のあり方
- 4) データの取り扱い 等

##### ② 具体的な評価の活用手法

費用対効果の評価結果を医療保険における医療技術の総合的な評価にどのように反映させるか

# 効果指標の取り扱い (その3)

福田参考人提出資料

# 目次

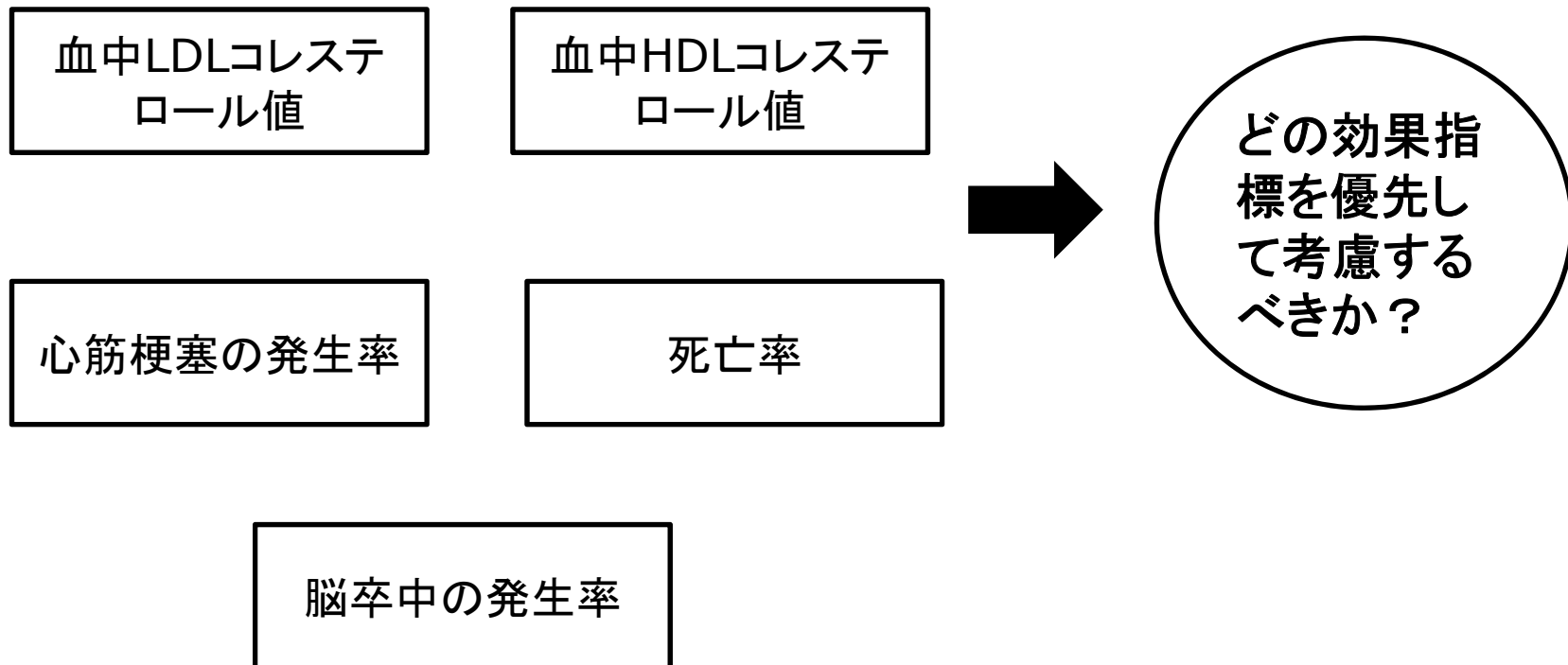
1. 効果指標の種類
2. 医師による評価と患者報告アウトカム
3. QALY(質調整生存年)
4. Appraisal(評価)

# 1. 効果指標の種類

# 医療技術の効果の多様性

- 個々の医療技術ごとに、得られる効果は異なる。
- 一定の普遍性・比較可能性等を有する評価指標を用いることが必要である。

・脂質異常症(コレステロール等の異常)の効果指標の例



# 医療におけるゴールの考え方(総論)

- 世界各国の有識者により、下記のコンセンサスがまとめられている。  
(現時点での世界的コンセンサス)

Goals of Medicine (1996)

(米国の生命倫理研究所(シンクタンク)の一つであるHastings Centerが、14カ国の有識者の意見を取りまとめて作成したコンセンサスレポート。)

## 医療におけるゴール

1. 疾病の予防と健康の増進および維持
2. 病気による苦痛の緩和
3. 病気を持つ者の治療とケア、および治療が不可能な者のケア
4. 若年期の死の回避と穏かな死の追求

# 医療におけるゴールの考え方(具体論)

- 前頁のGoals of Medicineを受け、より具体的に医療におけるゴールの考え方がまとめられている。

(スウェーデン ヨーテボリ大学) BENGT BRÜLDE *Health Care Analysis 9: 1-13, 2001.*

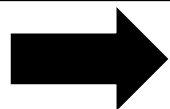
## The Goals of Medicine. Towards a Unified Theory

- 最も重要視されるもの

1. 寿命の保持・延長(特に、若年期の死を予防することによる。)
2. QoLの維持・向上(特に苦痛の除去による)

- 上記を達成するための医療的手段

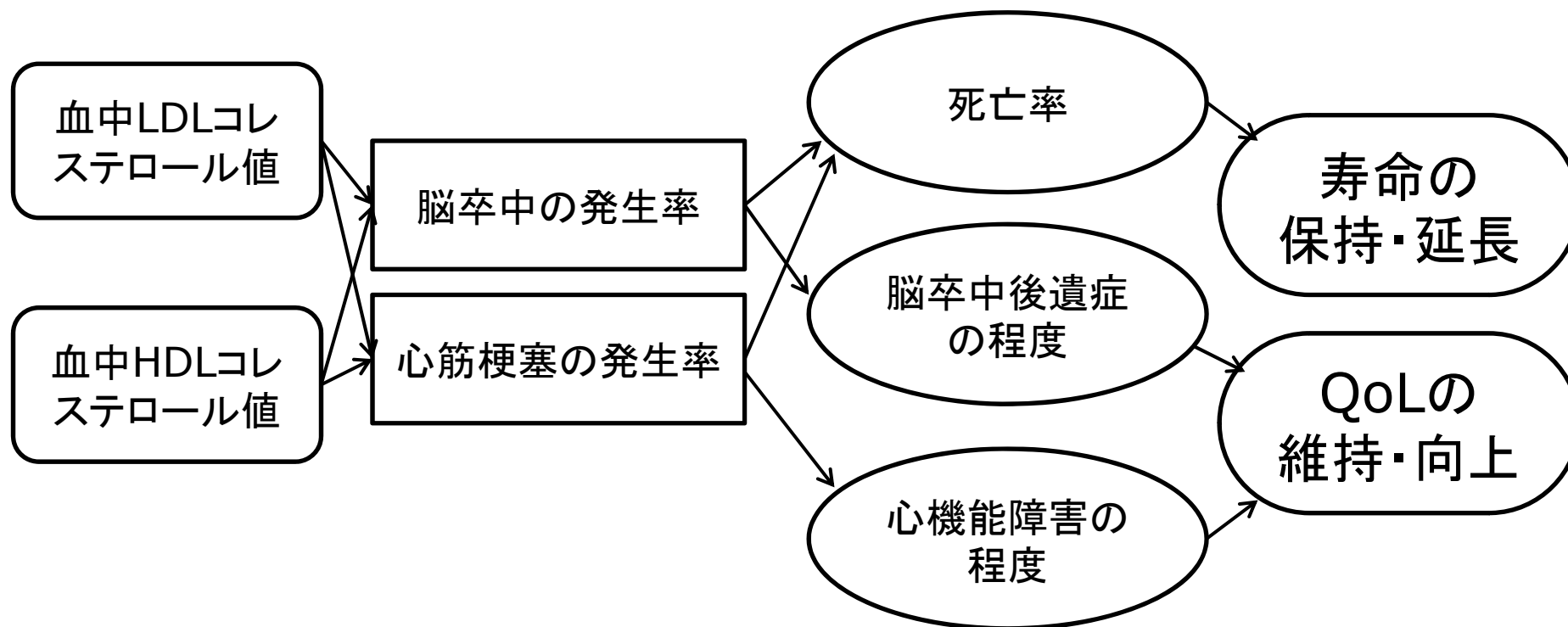
1. 健康関連の機能・能力の維持・向上
2. 構造的・機能的な健康状態の維持(特に疾病・負傷の予防及び治療による。)
3. (一定の症状・状態においては、)自分の症状・状態と調和できるようにすること。
4. (一定の症状・状態においては)疾病・負傷を予防するため、または自立した環境を維持・向上させるため、外的環境を向上させること。
5. 小児の成長・発達の促進



これらのゴールの達成度を測るための医療技術の  
効果指標はどのように設定するべきか？

# 医療技術の効果指標の分類イメージ (簡略化したもの)

各効果指標の「寿命の保持・延長」、「QoLの維持・向上」との関係  
(脂質異常症(コレステロール等の異常)の場合)





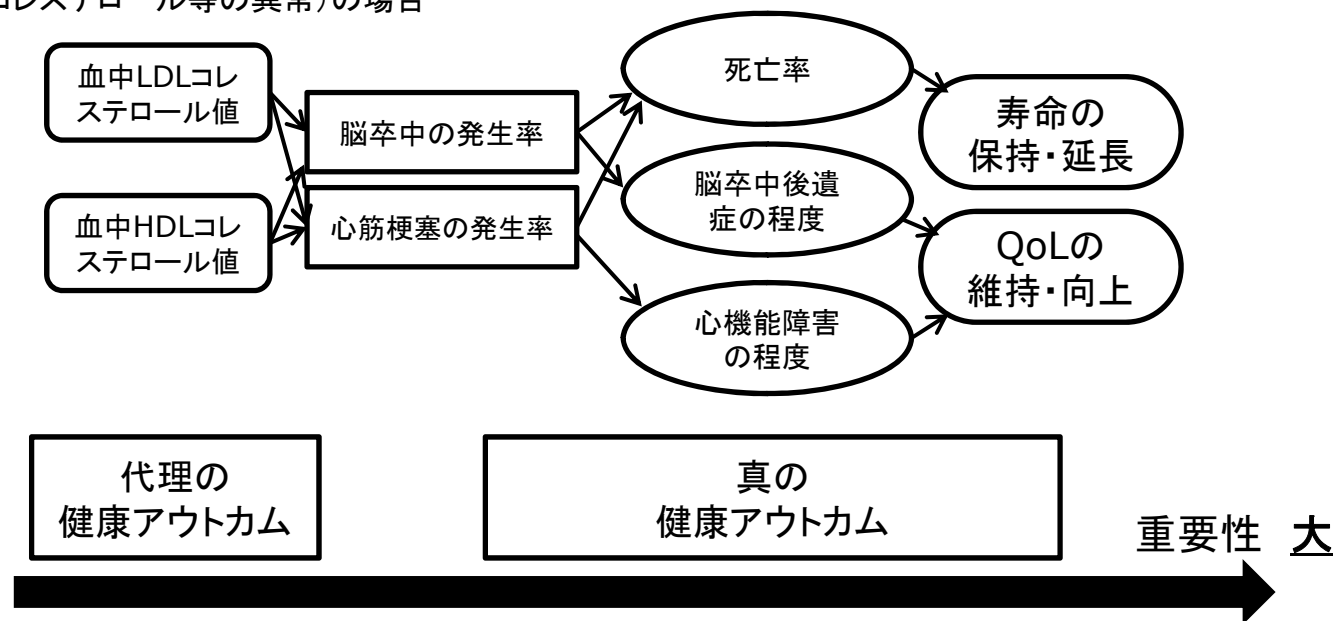
# 真の健康アウトカムと代理の健康アウトカム

- 健康アウトカムは例えば「真の(true)健康アウトカム」と「代理の(surrogate)健康アウトカム」に分類される。
- 真の健康アウトカムとは、目的に即したアウトカムである。  
(医療技術の効果を評価する目的の場合は、生存期間延長やQoL改善等)
  - 例えば、臨床試験の場合で、医療技術の評価を目的とせず、重大なイベント(心筋梗塞など)の発生率自体を比較する目的の場合は、これを真のアウトカムとする。
- 真の健康アウトカムの計測は、多くの被験者を長期間観察する必要がある等負担が大きい。
  - 真の健康アウトカムと相関があるとされる検査値等を用いることがある。代理の健康アウトカム

# 真の健康アウトカムとして用いられる指標と代理の健康アウトカムとして用いられる指標

- 真の健康アウトカムとして用いられることが多い指標  
生存期間、Quality of life (QoL)、主要なイベント(心筋梗塞等)発生率、治癒率など
- 代理の健康アウトカムとして用いられることが多い指標  
脂質異常症におけるLDLコレステロール値、糖尿病性合併症に対するHbA1c値、痛風の血清尿酸値、HIVにおけるCD4細胞数など

○脂質異常症(コレステロール等の異常)の場合



# 真の健康アウトカムを測定することの重要性

- 代理の健康アウトカムと真の健康アウトカムとが必ずしも正の相関を示しているとは限らないため、真の健康アウトカムを測定することが重要である。

## 【抗不整脈薬の臨床試験の例】

○死亡や心不全(真の健康アウトカム)の発症を減らすため、不整脈(代理の健康アウトカム)を減らす抗不整脈薬を投与したところ、逆に死亡や心不全が増加した。(真の健康アウトカムと代理の健康アウトカムの逆転)

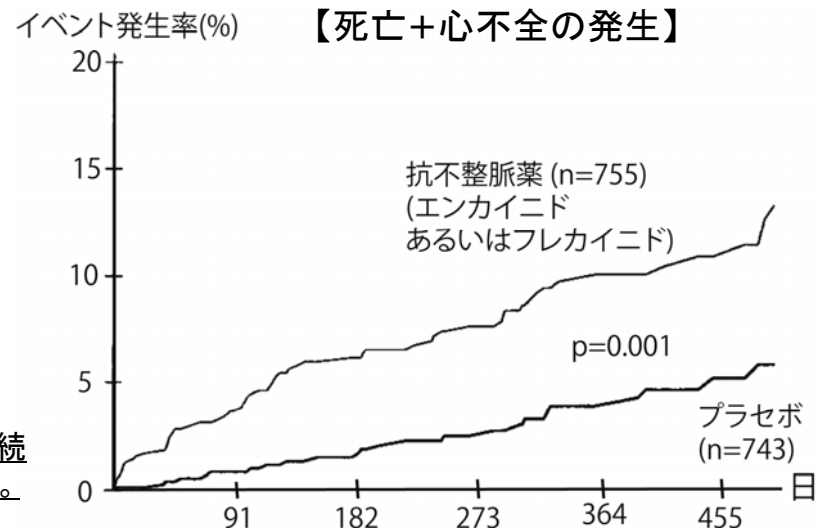
○対象:心室性不整脈のある心筋梗塞後の患者 (ランダム化比較試験)

抗不整脈薬(エンカイニドorフレカイニド)又はプラセボを投与し、死亡+心不全発症を比較。(抗不整脈薬で不整脈が減少することは、既に様々な臨床試験で確かめられている。)

### 【不整脈の発生】

プラセボ	多い
抗不整脈薬 (エンカイニドあるいは フレカイニド)	少ない

※心筋梗塞後で心室性不整脈(無症候性心室性期外収縮または非持続型心室頻拍)のある患者に対しては、当該薬剤は用いられていない。



# 生存期間を測定することの重要性

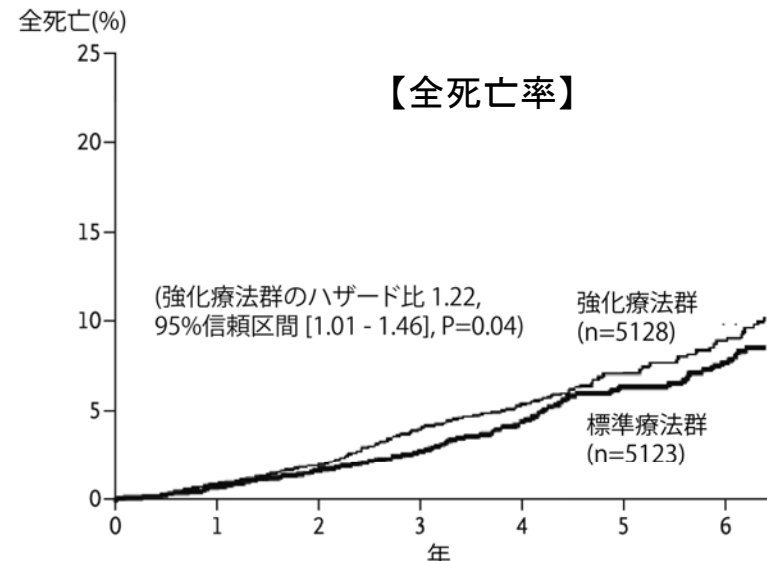
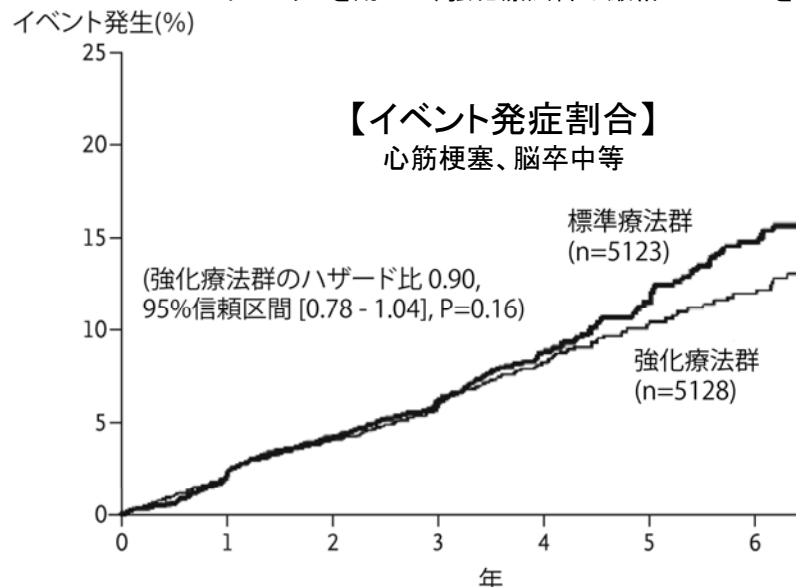
- 臨床試験の目的によっては、重大なイベント(心筋梗塞など)発生率等を真の健康アウトカムとすることも多い。
- しかし、重大なイベント発生率と生存期間やQoLとの間に必ずしも正の相関があるとは限らないため、生存期間やQoL自体を計測することが重要である。

## 【糖尿病治療の例】

○インスリンによる強化療法で、HbA1c(代理の健康アウトカム)を厳格にコントロールしたところ、重大なイベント発生率(心筋梗塞、脳卒中、心血管合併症による死亡)は減少傾向だったが、全死亡は増加してしまった。

○試験概要(対象: 2型糖尿病患者 (ランダム化比較試験))

インスリンを用いて、強化療法群は厳格にHbA1cをコントロール。標準療法群は、従来通りの基準でコントロール



The Action to Control Cardiovascular Risk in Diabetes Study (ACCORD)試験 (N Engl J Med. 2008;358(24):2545-59)

強化療法: 目標値HbA1c 6.0%以下  
標準療法: 目標値HbA1c 7.0~7.9%以下 **10**

# 健康アウトカム間にトレードオフがあるとき (肺癌の術後化学療法の例)(1)

【肺癌の術後化学療法の影響】

	化学療法あり	化学療法なし
癌の再発抑制効果	○ (あり)	× (なし)
副作用の発現	× (強い)	○ (なし)

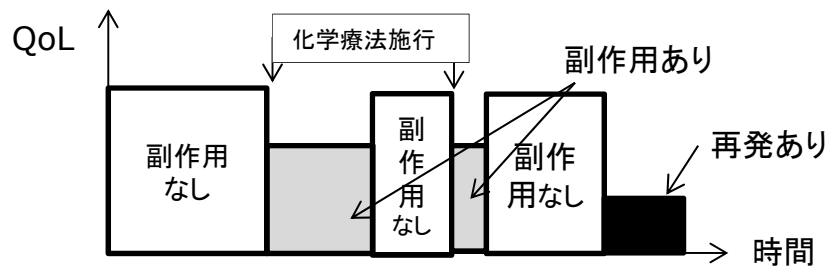
- ・術後化学療法の再発抑制効果が証明されている。
- ・カナダ国立がん研究所で行われた臨床試験

① 肺癌の再発抑制と術後化学療法の副作用の関係にトレードオフ(二律背反)がある。そのため、化学療法が有効であるにもかかわらず、患者が副作用を懸念して、なかなか普及しなかった。

(対象は低ステージ(stage Ib/ II)の非小細胞性肺癌)

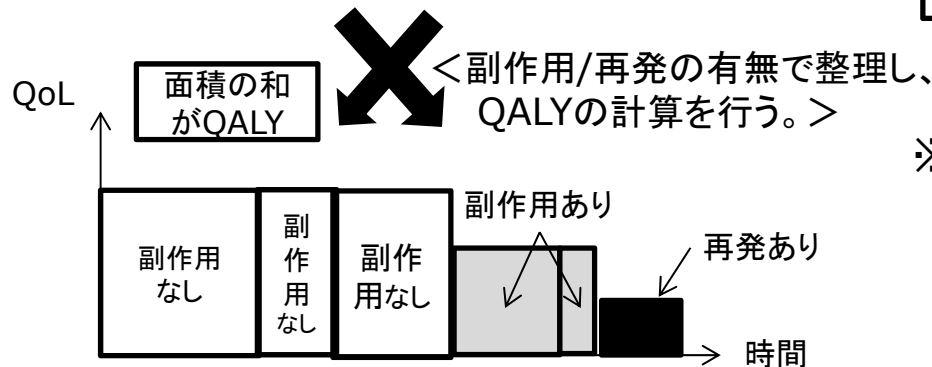
N Engl J Med 352:2589-2597, 2005

【術後化学療法を行ったときのQoL変化例(時系列順)】



② 癌の再発抑制効果と比べて、副作用によるQoLの低下がどれくらい大きいかを定量的に比較できれば、判断の補助として有用。

→ QALYを算出し判断の補助とした事例(左記、詳細次ページ)(J Clin Oncol. 2009;27(26):4268-73))

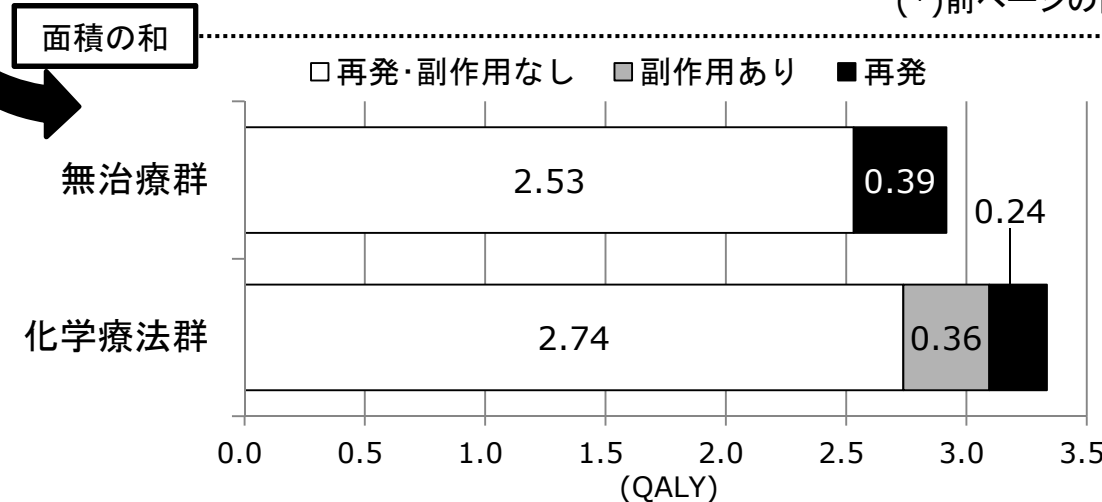
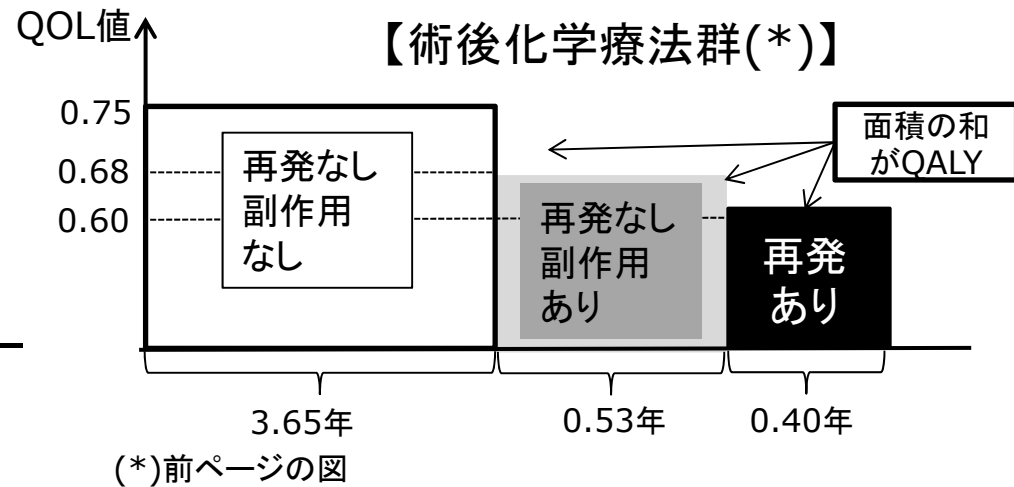
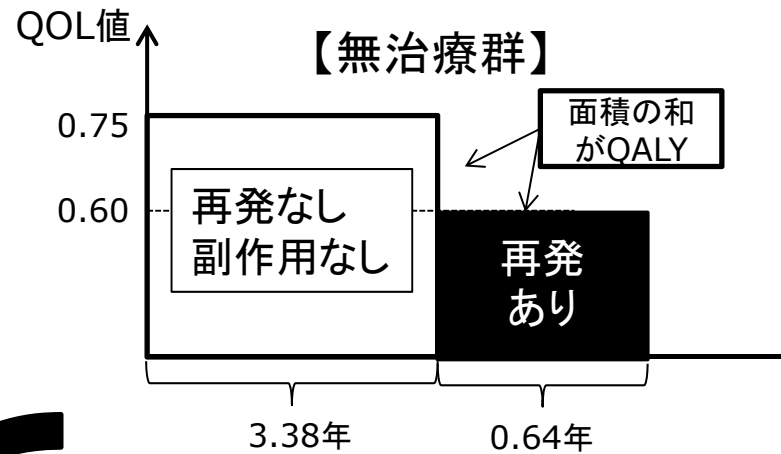


※ どちらの治療を選ぶかは、臨床的判断に加えて、患者自身が副作用や再発抑制効果をどのように評価するかも重要である。

詳細な数値等は次ページ参照

# 健康アウトカム間にトレードオフがあるとき (2)

- ・ 前頁における、肺癌術後化学療法におけるQALYの比較(カナダ国立がん研究所のデータ)
- ・ QALYの値は、下図における長方形(各健康状態を表している)の面積の和として計算できる



QALY=2.92

QALY=3.33

術後化学療法群で獲得できるQALYが無治療群より大きくなる。

## 効果指標の種類(まとめ)

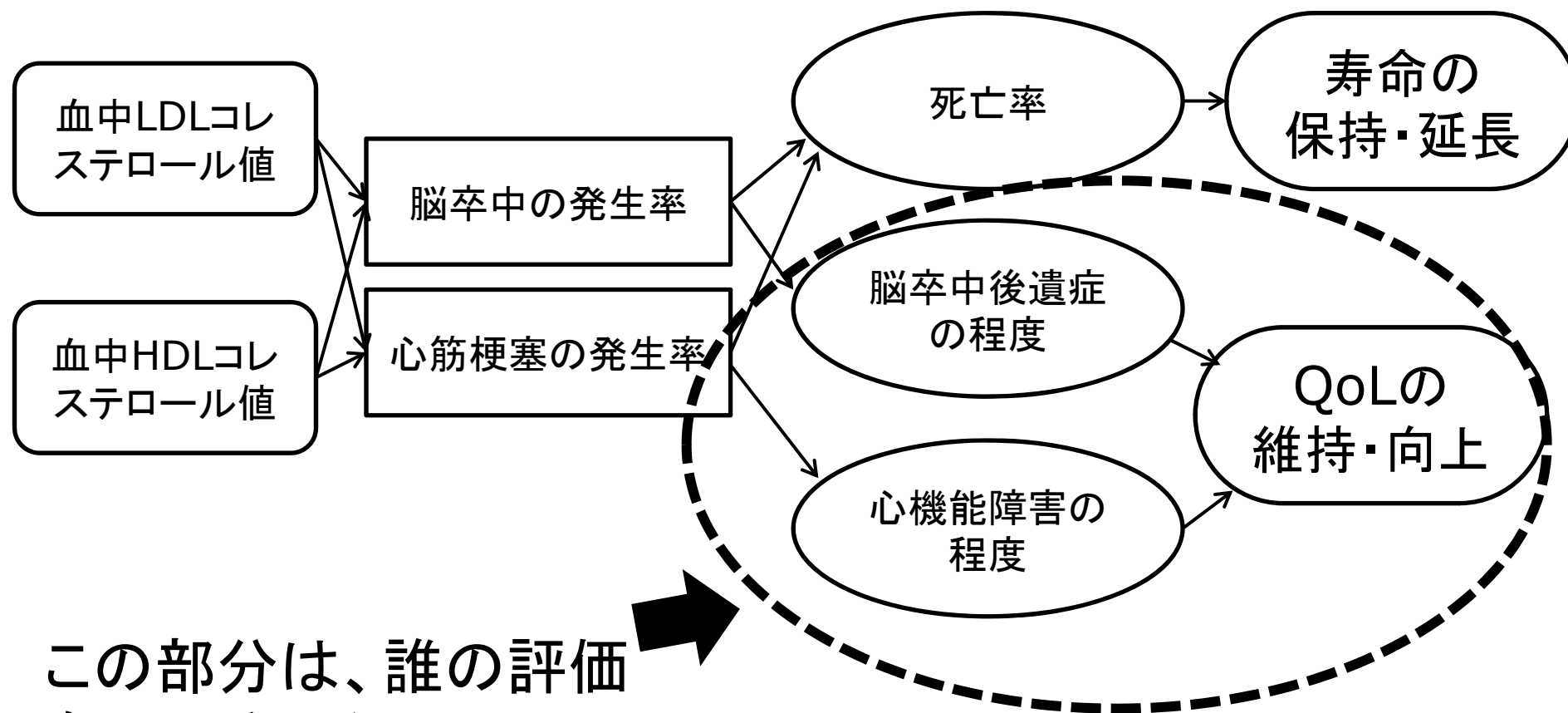
- 医療技術を評価するためには、代理の健康アウトカムより真の健康アウトカムが望ましい。
- 真の健康アウトカムの中でも、生存期間、QoLが重要である。(心筋梗塞などのイベント発生率は、生存期間やQoLと必ずしも正の相関を示さないことがある。)
- 複数の評価項目があり、それらの間にトレードオフがある場合、統合的な効果指標により、評価が容易になる。

## 2. 医師による評価と患者報告アウトカム



# QoL評価における「医師評価」と「患者報告アウトカム」

(脂質異常症に関連する医療技術の効果指標の分類イメージ(簡略化したもの))



この部分は、誰の評価を用いるのか？

# 医師による評価指標の例

○ 患者の状態を医師が客観的に評価し、指標化する。

【修正Rankinスケール(mRS)】  
脳卒中後の後遺症の程度を  
医師が基準に沿って評価

【NYHA分類】  
心機能障害の程度を  
医師が基準に沿って評価

## 日本版modified Rankin Scale(mRS)

0	まったく症候がない
1	症候はあっても明らかな障害はない: 日常の勤めや活動は行える
2	軽度の障害: 発症以前の活動がすべて行えるわけではないが、自 分の身の回りのことは介助なしに行える
3	中等度の障害: 何らかの介助を必要とするが、歩行は介助なしに行 える
4	中等度から重度の障害: 歩行や身体的要求には介助が必要である
5	重度の障害: 寝たきり、失禁状態、常に介護と見守りを必要とする
6	死亡

## New York Heart Association(NYHA)分類

I 度	心疾患はあるが身体活動に制限はない。 日常的な身体活動では著しい疲労、動悸、呼吸困 難あるいは狭心痛を生じない。
II 度	軽度の身体活動の制限がある。安静時には無症状。 日常的な身体活動で疲労、動悸、呼吸困難あるい は狭心痛を生じる
III 度	高度な身体活動の制限がある。安静時には無症状。 日常的な身体活動以下の労作で疲労、動悸、呼吸 困難あるいは狭心痛を生じる
IV 度	心疾患のためいかなる身体活動も制限される。 心不全症状や狭心痛が安静時にも存在する。わず かな労作でこれらの症状は増悪する。

# 医師による評価と患者報告アウトカムの関連について (脳卒中後遺症の評価指標の例)

- ・ mRS(医師による評価;前頁参照)とQoL尺度(EQ-5D\*;患者による評価)の関係性を大規模に調査。回帰分析等を行い、両尺度の関連を解析した。
- ・ mRSの各段階がEQ-5Dスコアでどの程度となるかを定量化している。

【mRSとEQ-5Dスコアの対応】(発症後24ヶ月の値) n=425

mRS	0	1	2	3	4	5	6
EQ-5D	0.959	0.812	0.656	0.545	0.248	0.020	0.0

調査は英国で実施

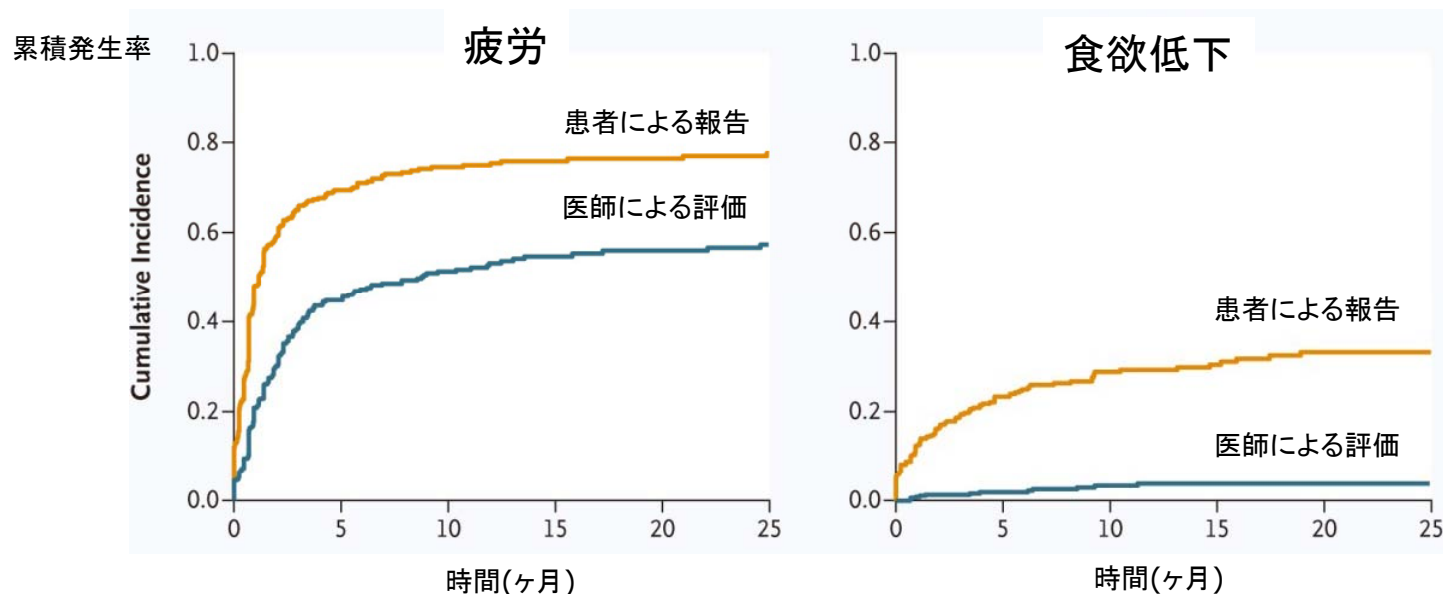
※)EQ-5Dとは  
健康関連QoLを測定するために開発された包括的な評価尺度  
1987年に設立されたEuroQoL グループが開発  
「完全な健康=1」「死亡=0」と基準化された健康状態のスコアを算出

Rivero-Arias O, Ouellet M, Gray A, et al. Mapping the modified Rankin scale (mRS) measurement into the generic EuroQoL (EQ-5D) health outcome. Med Decis Making. 2010;30(3):341-54.

# 「医師の評価」と「患者報告アウトカム」の差

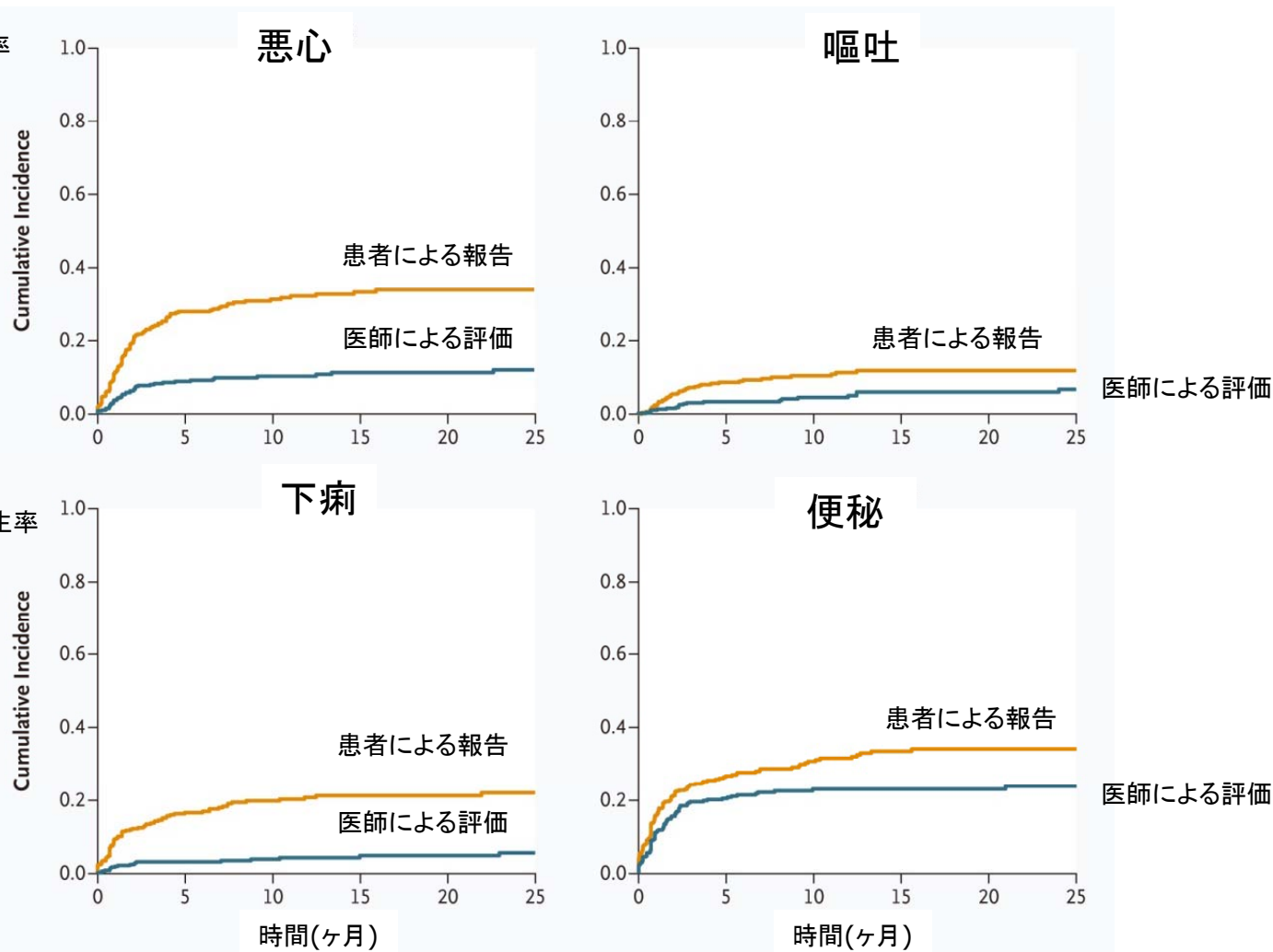
- QALYで使用するQoLは「患者報告アウトカム(Patient Reported Outcome: PRO)」の一種である。
- 「医師による評価」と「患者の報告」との間には正の相関があるが、その値は異なることがある。

(例)467人の癌患者に対して、抗癌剤の副作用を医師による評価(CTCAEに基づく)と患者による報告で比較したもの。(N Engl J Med. 2010;362(10):865-9)



(続き)

累積発生率



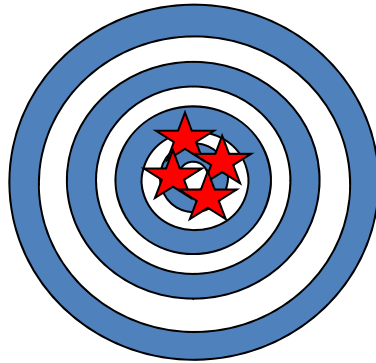
■「医師による評価」は「患者による報告」よりも副作用を小さく評価しがちであった。

→ 副作用が減少するような医療技術の効果を小さく評価してしまう可能性がある。

# 患者報告アウトカム(PRO)の精緻化

- QoLなどの患者報告アウトカム(PRO)として用いられる尺度は、単なるアンケートとは異なる。
- 患者の報告を定量化し、科学的に評価するため、以下のような計量心理学的特性が検討された上で、尺度として確立してきた。
  - 信頼性 Reliability:再現可能で一貫した結果が得られているか
    - ・ 再現性(繰り返し測定しても同じ値となるか)
    - ・ 内的整合性(項目に重複がないか)
  - 妥当性 Validity:測定したいものが測れているか
    - ・ 内容的なもの(的確性、網羅性)
    - ・ 基準関連的なもの(既存の基準との整合)
  - 応答性 Responsiveness:測定したいものの変化を検出できるか
  - 実施可能性 Feasibility:調査の実施が現実に行えるか

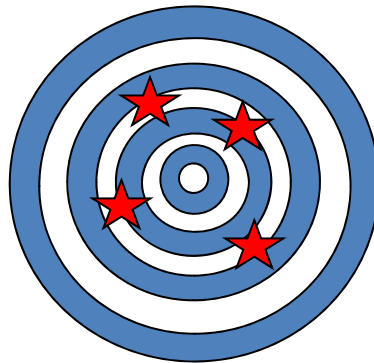
# 測定の信頼性と妥当性(イメージ)



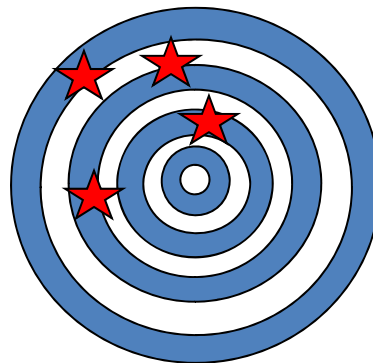
信頼性は**高い**  
妥当性は**高い**



信頼性は**高い**  
妥当性は**低い**



信頼性は**低い**  
妥当性は**高い**



信頼性は**低い**  
妥当性は**低い**

■ 信頼性: 矢の当たっている位置 (=個々の測定値)はばらついていないか?

■ 妥当性: 的の中心 (=測定したいもの)に矢が当たっているか?

信頼性・妥当性がともに高くなる (図左上)ように尺度が開発される。

(mRSとEQ-5Dの関係性も基準関連妥当性の一例)

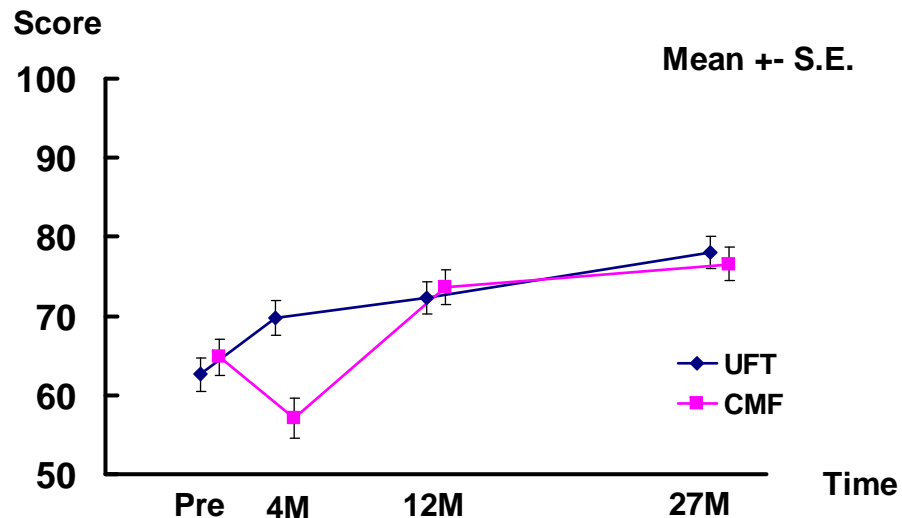
# 患者報告アウトカム(PRO)の検証の例 (応答性)

- QoL尺度などの患者報告アウトカム(PRO)の信頼性、妥当性、応答性等は、定量的な調査を実施後、統計学的見地等から検証されている。

## 【応答性の検証の例】

- ・乳癌患者を対象
- ・化学療法中とその後のQoLを調査したもの。
- ・経口剤(UFT [テガフル・ウラシル])と注射剤(CMF[シクロホスファミド、メトトレキサート、フルオロウルシルの3剤併用療法])の比較を行った。
- ・患者報告アウトカム指標: EORTC(European Organisation for Research and Treatment of Cancer: 欧州がん研究・治療機構)が、癌患者対象に開発したQoL尺度

## EORTC < Global quality of life >



化学療法期間中(4ヶ月目)で経口剤のQoLが注射剤より有意に大きくなっている。

→ 反応すべきところでQoLが反応している。

ANOVA P=0.025



# 米国における「患者報告アウトカム」の取り扱い

- 2009年12月に、FDA(米連邦食品医薬品庁)は、患者報告アウトカム(PRO)に関する指針を発行し、以後、医薬品・医療機器の申請における使用を認めることとした。

「企業向け指針 患者報告アウトカムの測定法について/医薬品・医療機器における承認申請のための方法」(抜粋) FDA(米連邦食品医薬品庁)

- 一般に、適切にデザインされた研究で明確に定義された信頼性の高い PRO 尺度で測定された所見は、適応申請の範囲が、尺度の測定能力と矛盾がない場合、医薬品/医療機器の承認申請に使用することができる。
- 患者が最もよく知っている、患者の視点から測定することが最も適しているような概念の測定においては、PRO 尺度を使うことが推奨される。

(国際医薬経済・アウトカム研究学会日本部会による翻訳)

# 医師による評価と患者報告アウトカム(まとめ)

- 患者報告アウトカム(PRO)は、定量的な調査、統計学的解析等により検証され、精緻化されてきている。
- PROの中でも、特にQoL尺度は、患者自身の医療技術に対する価値付けを加味できる等の利点があり、医療技術の効果をより幅広く捉えるための指標として重要である。
- QoL尺度などのPROについては、医師の評価との対応を明らかにする研究が実施されている。両者は全く異なるものではなく、関係性が大きいことが明らかになっているが、PROの方が、患者に生じた変化を鋭敏に捉えられると言われている。

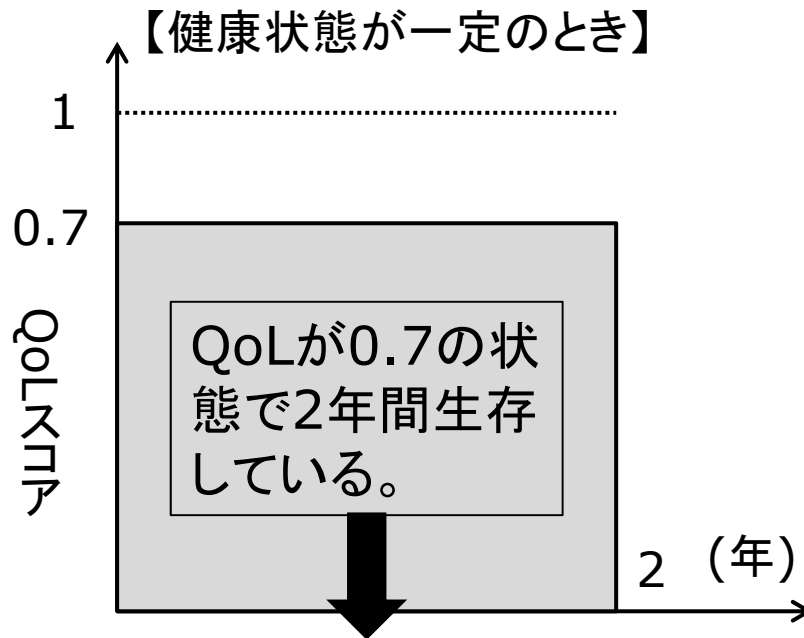
(参考) J Natl Cancer Inst. 2011;103(24):1808-10

- 患者自身の「主観的」な経験や治療に対する価値づけは、その他の誰にも知ることができない。
- もちろん、医師による「客観的」な評価も重要であり、それらを組み合わせることにより、より有益な情報が得られる。

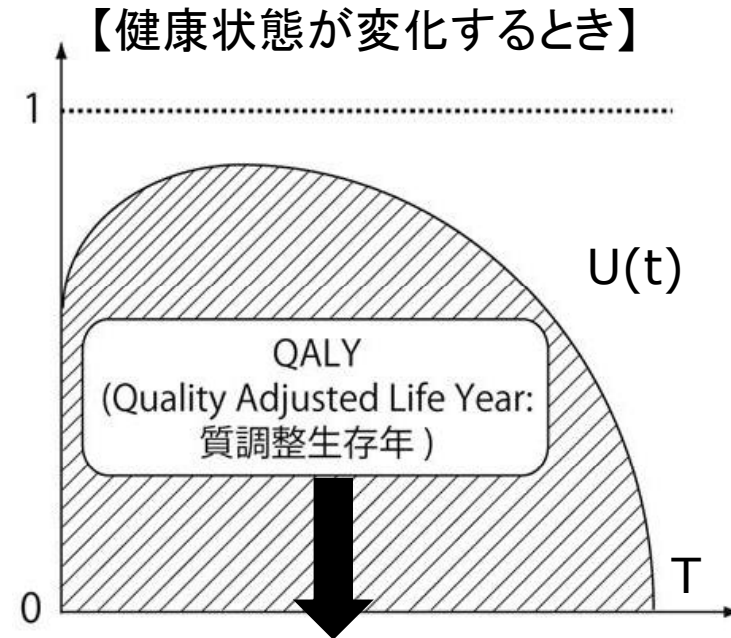
### **3. QALY (質調整生存年)**

# QALY(質調整生存年)の計算方法

- 生存年数と生活の質(QoL)の双方を考慮する。
- QoLについては、1を完全な健康、0を死亡とする「QoLスコア(効用値)」を用いる。
  - ある健康状態でのQALY = 【QoLスコア】 × 【生存年数】



「 $QALY = 0.7 \times 2 = 1.4$ 」と計算できる。



曲線下の面積で計算できる。

$$QALY = \int_0^T u(t) dt$$

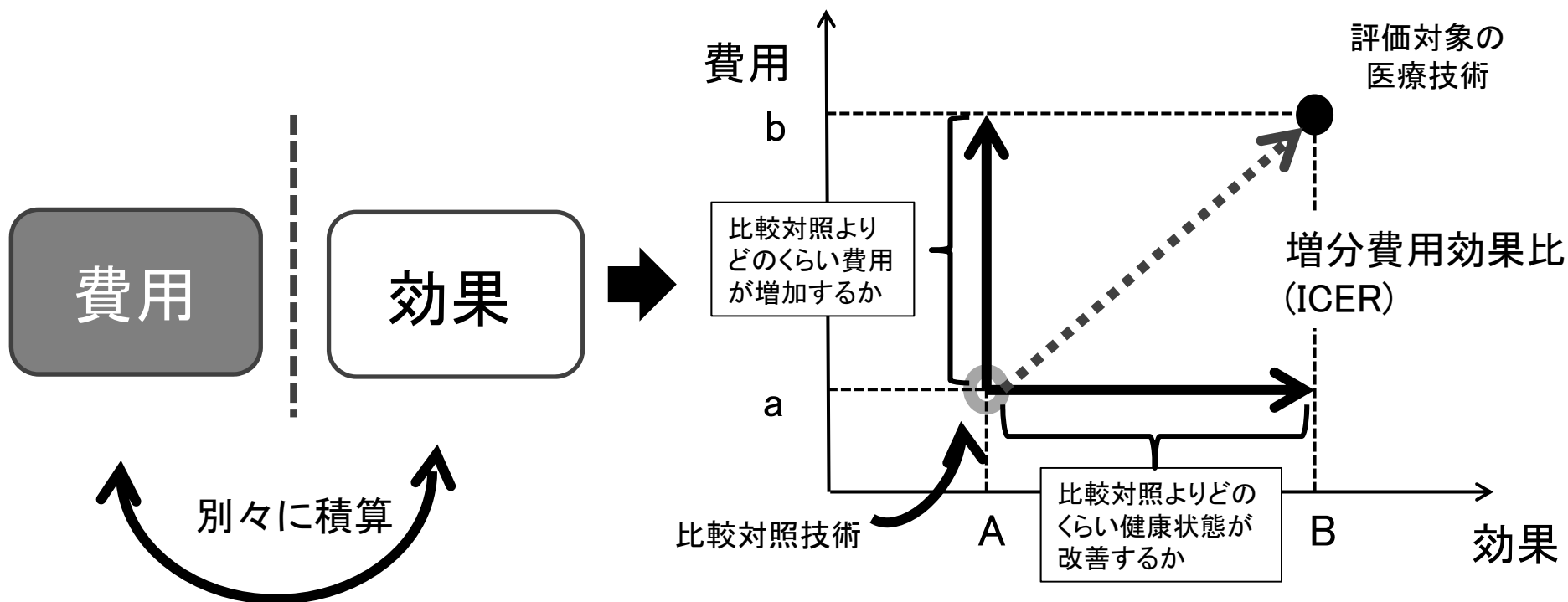
# (参考) 費用効果分析の手順

(1) 評価対象の医療技術及び比較対照の医療技術の「費用」と「効果」を別々に積算する。

(効果は金銭化しない)

(2) 評価対象の医療技術と比較対照の医療技術との増分費用効果比の評価を行う。

$$\text{増分費用効果比 (ICER)} = \frac{b-a \text{ (費用がどのくらい増加するか)}}{B-A \text{ (効果がどのくらい増加するか)}}$$



# 各国のガイドラインにおけるQALYの取り扱い

---

第1選択としてQALYを推奨

イギリス、アイルランド、ノルウェー、  
ニュージーランドなど

---

QALYも含めて推奨

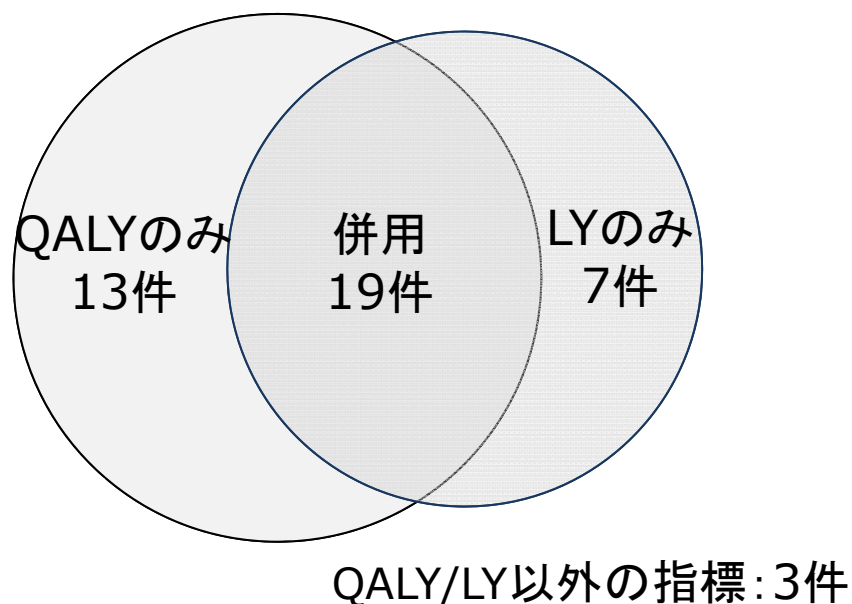
オーストラリア、カナダ、スウェーデン、  
オランダ、ベルギー、フィンランド、  
韓国、台湾など

---

費用対効果評価を意思決定に活用している国で、QALYを選択肢として排除している国は、ほぼ存在しない。

# (参考) QALYが必須ではない国の状況 (オランダの例)

- オランダでは、QALYを用いた評価が必須ではないが、実際の評価は、QALYを用いている例が多い。
- 費用対効果を計算している42件の内訳
  - 32件は、QALY効果指標に使用
  - 7件は、QALYは使用しないがLY(生存年)を使用。



## 【内訳】

- QALY (質調整生存年)のみ使用の例 (13件)
  - 生活習慣病(糖尿病など)
  - 眼科疾患(緑内障・加齢性黄斑変性症)
  - 関節リウマチ治療薬
- LY (生存年)のみ使用の例 (7件)
  - 侵襲性カンジダ症治療薬
  - 一部のがんに対する治療薬
- QALY とLYの双方を使う例 (19件)
  - 抗がん剤
  - 抗血小板薬
- QALYもLYも使わない例 (3件)
  - 皮膚科疾患治療薬
  - 麻薬性鎮痛薬
  - 放射線増感剤

## QALYが費用対効果評価において汎用される理由(1)

### (a) 多くの疾患においてQALYを用いて評価できる。

- 例えば、逆流性食道炎における胸焼けの改善等(特異的指標:GERD-HRQL- Questionnaire等)と過敏性腸症候群における腹痛の改善等(特異的指標:IBS-SSS等)とは、いずれもQoLによる評価であるが、指標が異なり、直接は比較できない。  
→ 逆流性食道炎も過敏性腸症候群も同じQALYで評価可能。
- 多くの疾患で共通の効果指標としては、生存期間がある。しかし、生存期間の改善が少なく、QoL改善が重要な場合(例えば関節リウマチ、過敏性腸症候群など)は、生存期間を指標としても、医療技術の価値が十分に評価できない。  
→ QALYであれば、生存期間が変化せずとも、QoLが改善する医療技術の価値を評価可能。



## QALYが費用対効果評価において汎用される理由（2）

### (b) 複数の効果を同時に評価できる

- 複数の重要な効果指標(例えば、糖尿病で「HbA1Cの低下」と「低血糖発作頻度の低下」)がある場合、それぞれの指標ごとに複数の費用対効果評価を行っても、その医療技術を総体として評価することが困難である。
- 効果指標間に何らかのトレードオフ(有効性は大きいが、副作用も大きいなど)がある場合、複数の指標を用いる評価の解釈はさらに難しくなる。
- QALYであれば両者の関係がQoLを通して一つの尺度となるため、その医療技術を総体として取り扱うことができる。

【スライド11及び12 「健康アウトカムにトレードオフがあるとき」参照】

## QALYが費用対効果評価において汎用される理由 (3)

### (c) 結果の解釈がしやすい

- 例えば、「C型肝炎ウイルスの除去率が10%改善する」といわれても解釈が難しい。

→ QALYに換算すると、判断がしやすい。

(イギリスNICEにおける例)

- C型肝炎治療における、「ペグインターフェロン $\alpha$ /リバビリン併用療法」への「ボセプレビル」の上乗せ治療の評価
  - 「ペグインターフェロン $\alpha$  /リバビリン併用療法」のみと比較して、1QALY獲得するのに£11,601かかると推計された。(未治療の患者を対象とする場合)
  - これは、NICEの基準(£20,000 ~ £30,000/QALY)に照らして、費用対効果がよいとされた。

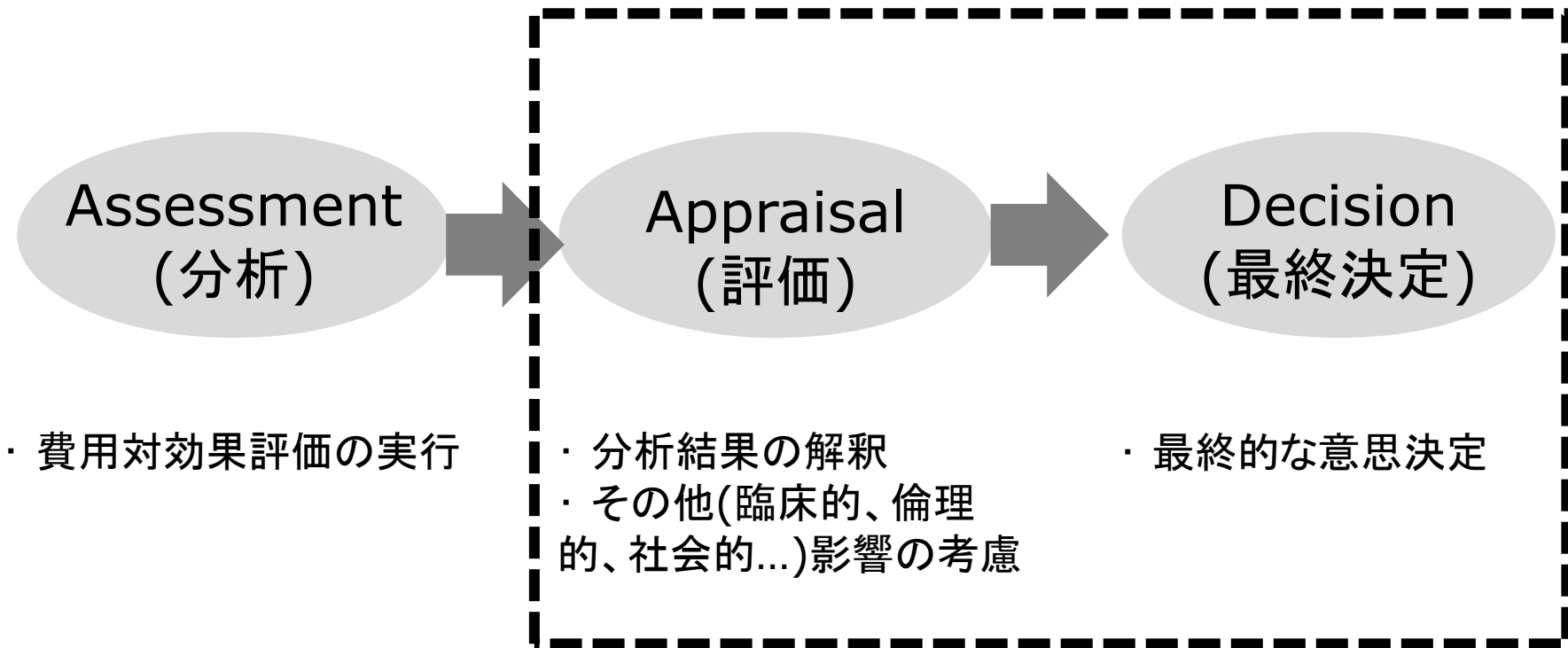
# QALY(質調整生存年)(まとめ)

- QALYには、医療技術の効果指標として以下のような利点がある。
  - 多くの疾患で使用できる。
  - 複数の効果を同時に評価できる。  
(トレードオフがあっても、一つの指標に統合して判断が可能)
  - 結果の解釈がしやすい。

## **4. APPRAISAL (評価)**

# Appraisalについて

一般的な費用対効果評価を用いた意思決定のプロセス



Appraisalにおいては、様々な要素を考慮して評価を行う。

# Appraisalで考慮される 増分費用効果比以外の要素（1） イギリスの例

- NICEが(QALYを用いた)費用対効果以外にも定性的に勘案する要素の例
  - Severity: 疾患の重症度
  - End-of-life: 致死的な疾患での延命治療
  - Stakeholder persuasion: 関係者からの意見
  - Significant innovation: イノベーションの大きさ
  - Disadvantaged population: 障害者などの集団への配慮
  - Children: 小児の疾患

# イギリスにおけるappraisalの具体例

小児骨肉腫に対するMifamurtideの評価例 (TA235) 2011年

1. 当初の分析結果は、費用対効果が劣るとされた。(£56,700/QALY)
  2. 下記の点を変更し、再度分析を実施した。
    - 小児であることを考慮して、長期的な予後を重視するような分析方法とする。
    - 価格を割り引く。→結果、費用対効果は改善したが、まだ課題があるとされた。(£36,000/QALY)
  3. NICEは以下の点を考慮して、Mifamurtideを推奨した。
    - Mifamurtideが革新的であること。
    - 稀少疾患に対する治療薬であること。
    - 家族等のQoLの損失が費用対効果に反映されていないこと。
- このように、費用対効果評価の結果を機械的に用いているわけではなく、Appraisalのプロセスで様々な観点を考慮する。

※) NICEでは1QALYあたり£20,000~£30,000を費用対効果が優れる目安としている。

# Appraisalで考慮される 増分費用効果比以外の要素 (2) スウェーデンの例

- スウェーデンにおける医薬品償還3原則

(a) "the human value principle"

すべての人間が平等に取り扱われる、年齢や人種、性別等による差別を受けない。

(b) "the need and solidarity principle"

重症度の高い人は優先的に償還される

(c) "the cost-effectiveness principle"

費用対効果のよい医薬品を償還する



# Appraisal(評価)(まとめ)

- 医療技術の評価のプロセスには、「Appraisal(評価)」を設けることが一般的
- Appraisalにおいては、下記のような様々な要素を考慮して評価を行う。
  - 臨床的背景
  - 倫理的背景
  - 社会的背景

# まとめ

- 様々な効果指標の特徴を踏まえて、評価に用いることが重要である。
- QALYは、その包括性や比較可能性が高いなど費用対効果評価を行う上での利点が多いため、医療技術に共通の尺度として、QALYを含めた評価を行なうことは重要である。
  - 一方、その他の指標を用いた解析も、医療技術の特性に応じて考慮する必要がある
- 効果指標を定めて評価を行っても 機械的に意思決定に用いるのではなく、その分析結果をどのように解釈するか「評価(appraisal)」のプロセスが重要である。