

簡便な効用値算出法の開発：日英国際比較研究

山本 洋介

京都大学医学部附属病院臨床研究総合センター／京都大学大学院医学研究科医療疫学分野 講師
(助成時：京都大学大学院医学研究科医療疫学分野 講師)

今回は、ファイザーヘルスリサーチ振興財団に助成していただき、また、このような発表の機会を与えていただきまして、誠にありがとうございます。

【スライド-1】

私は、簡便な効用値算出法に関する研究を行いましたので、それに関する発表をさせていただきます。

【スライド-2】

まず、「効用値とは」ということですが、皆さんご存じと思いますが、健康状態の価値付けを可能にするQOL尺度の一つでutilityとも申します。これ自体はQOLを一次元の概念として、ゼロから1の間の数値として表します。原則として死亡をゼロ、完全な健康状態を1として患者さんの健康状態を評価するということになっていますが、例えば、糖尿病性網膜症に伴う失明による健康状態を、例えば0.7などと表すことができ、さまざまな疾患間でのQOLの状態を一つの物差しに乗せることが可能になります。これによって、費用効果分析など医療の経済評価が容易に可能になり、近年その分野で多用されています。

なお、効用値に関しては、今まで、standard gambling法とか、あるいはtime trade-off法といった方法で算出されることが多かったのですが、より簡便に算出する方法として、尺度から推定するという方法が近年しばしば用いられています。それにはSF-36から推定できるSF-6D、あるいは5項目からなるEQ-5D、さらにはHUI(health utility index)などの尺度が含まれます。

スライド-1

簡便な効用値算出法の開発に関する研究

2014.11.29

京都大学大学院 医学研究科 医療疫学分野
京都大学医学部附属病院 臨床研究総合センター
山本 洋介

スライド-2

背景：効用値とは

- 健康状態への価値づけを可能とする
 - QOLを一次元の概念とし0-1の数値で表す.
 - 原則として、死亡=0, 完全な健康状態=1とし、患者の健康状態を評価.
(例)糖尿病性網膜症による失明 =0.69 など
 - 費用効果分析など医療の経済評価で多用
- 尺度から効用値を推定・算出する方法
 - SF-6D(SF-36に基づき推定)
 - EuroQol (EQ-5D)
 - Health Utility Index (HUI)...など

【スライド-3】

日本語版SF-6Dに関しては、すでに開発と検証が終わってしまっていて、6つの下位尺度からなる尺度として利用が始まっています。

その6つの下位尺度とは、ここに挙げている身体機能、日常役割機能、社会生活機能、身体の痛み、心の健康、活力といったもので、これらの組み合わせで1万8,000通りの健康状態の表現が可能となっています。オリジナルに則した翻訳にはなっていますが、日本においては身体・精神に関する日常役割機能は同じ回答パターンを示すと言われるため、全体的な頻度を問う質問に変更されたという経緯があります。

【スライド-4】

しかしながら、このSF-6Dの運用上の問題点として、効用値算出のために現在この36項目全て回答することが求められているという点が挙げられます。これ自体は非常に項目数が多く、効用値を求めるだけのためには効率が悪いということが挙げられています。そのため、より少ない項目で健康関連QOLを測定するSF-8やSF-12から効用値が求められれば、非常に簡便なツールとなり得ることが期待されています。また、今まで蓄積されてきたSF-8やSF-12のデータを用いた効用値の算出が可能となり、既存のデータを用いた費用効果分析などの研究の発展が期待できると思われまます。

【スライド-5】

そこで、私たちはMapping Approachというものをを用いてSF-8やSF-12の回答から効用値の推定を試みしたので報告します。これは簡単に説明しますと、まず、SF-36ならびにSF-8

スライド-3

3

日本語版SF-6Dの概要

- 6つの下位尺度
 - 身体機能 6段階
 - 日常役割機能 5段階
 - 社会生活機能 5段階
 - 身体の痛み 6段階
 - 心の健康 5段階
 - 活力 5段階

6項目の組み合わせで、18000の異なった健康状態の表現が可能
111111から655655まで

<オリジナル版との違い>
日本においては、身体・精神に関する『日常役割機能』は同じ回答パターンを示すため、全般的な頻度を問う質問に変更。

スライド-4

4

SF-6Dの現在の問題点

- SF-6Dの効用値算出のためには、現在SF-36の全て回答が原則
 - ⇒ 項目数が多く、効用値を求めるだけのためには効率が悪い
- SF-8・SF-12から求められれば、非常に簡便なツールとなりうる。
 - ⇒ 既存のデータから効用値の算出が可能に
 - ⇒ Mappingの手法を用いる

スライド-5

5

Mapping Approach

The diagram shows two boxes on the left: 'SF-8/12の回答項目' and 'SF-36の回答項目'. Arrows from both boxes point to a box on the right labeled '効用値'. A curved arrow labeled 'MAPPING' connects the two boxes. Below the boxes, text reads: 'NORM2007データを derivation・validation用にランダムに分割' and '既にSF-6Dによって効用値の算出が可能'.

がともに含まれた質問紙調査のデータを用います。第一ステップとして、SF-36の回答項目を用いて効用値を算出します。これは先ほど申した既存のアルゴリズムで換算することができます。これを従属変数として、それに対してSF-8やSF-12の回答項目を説明変数とするような当てはまりの良いモデルを構築するという、原理としては非常に単純な仕組みになっています。

【スライド-6】

今このことをもう一度かい摘んで説明しますと、SF-36の結果から効用値を算出します。ただ、日本で今までマッピングを用いた効用値の算出がなされたことがないので、英国の結果と比較して、日本のデータのマッピングの可能性を検討しました。

その上で、推定された効用値を従属変数、SF-8ならびにSF-12の項目を説明変数とした複数のモデルを作成して、当てはまりの評価を行いました。

さらには、当てはまりのよいとされたモデルを別の validation sample を用いた検証を実施して確認するというプロセスを含んでいます。これはSF-12においても同様のことをしていますが、今回はSF-8の結果についてのみご説明します。

スライド-6

6

方法: Mapping Approachの概要

1. SF-36の結果から効用値を算出, 英国の結果と比較してMappingの可否を検討
2. 推定された効用値を従属変数, SF-8の項目を説明変数とした複数の回帰モデルを作成. 当てはまりの良さからモデルを選択
3. 推定されたモデルの回帰式を用いて, 別の validation サンプルでモデルの検証を実施.
4. SF-12においても1~3のプロセスを実施(今回は時間の都合上割愛します)

【スライド-7】

初期モデルの作成に関しては、まず、以下の2つの観点から作成し比較しました。

まずは、説明変数の候補を何にするかということですが、1つ目のモデルは、メンタル・コンポーネント・サマリー (MCS)、フィジカル・コンポーネント・サマリー (PCS)、ロール・コンポーネント・サマリー (RCS) を説明変数の候補としました。これらはSFツールの回答結果から算出される身体・精神・役割機能の各サマリースコアを連続変数として用いたモデルです。2つ目は8つの下位尺度得点を連続変数として用いたモデル。3つ目は、この8つの下位尺度の結果を順序変数として扱い、ダミー変数化して用いたモデル、以上のような説明変数の候補となるセットを準備しました。

なお、実際に適用する初期モデルとしては、最尤推定を用いた線形回帰分析と、あとはTobitモデルの2つ

スライド-7

7

方法: モデル作成(初期モデル)

■ 以下の2点の観点から各モデルを比較.

- ① 説明変数の候補(従属変数は効用値)
 - **Model 1:** MCS, PCS, and RCS scores as continuous variables
 - **Model 2:** the eight SF-8 dimension scores as continuous variables
 - **Model 3:** the dummies of each SF-8 item
- ② 適用するモデル
 - maximum likelihood linear regression (ML)
 - **Tobit model**

をモデルとしています。

【スライド-8】

そして、この初期モデルのほうで最も当てはまりがよいとされたモデルを、またさらに当てはまるように拡張していくことをしています。例えば説明変数に対してMOST term、これはSF-8の項目で最悪の項目を付けたような人を示す変数ですが、こういうものを説明変数に加えました。さらには性、年齢などを加えて、当てはまりを改善するという方法を試みました。

そしてさらには、2段階モデルへの応用も試みました。具体的には、1つ目の最尤推定法に基づく線形回帰を行った後に、リニア・スプラインモデルあるいはキュービック・スプラインモデルを用いました。

【スライド-9】

モデルの当てはまりの評価に関しては、平均・標準偏差、最小値・最大値、二乗平均平方根誤差、平均絶対誤差、平均誤差、さらには赤池情報量基準など、多面的に判断を行っています。また、今回の発表では割愛していますが、従属変数の効用値を区間ごとに分け、その各区間での当てはまりの評価も実施しています。

【スライド-10】

さて、結果ですが、対象者の特性を述べさせていただきます。

2,304名の患者さんに対して、2対1で Derivation data set と Validation data set に分割しています。その患者さんの背景は年齢51歳が平均で、女性が48～49パーセント。SF-6Dは0.68という平均値でした。

そして、分布に関しては右の図で

スライド-8

8

方法：初期モデルの拡張

- 初期モデルで良好な結果を得られたモデルを拡張し当てはまりを改善

説明変数

- Model 4: the dummies of each SF-8 item + MOST terms
- Model 5: the dummies of each SF-8 item + age, gender, MOST terms

※MOST terms: a dummy representing a severest health state in any items

- 2段階モデルを利用 (step1: ML, step2: linear spline)

スライド-9

9

方法：モデルの当てはまり評価

- 以下の指標から判断(総合的に判断)

- Mean/SD
- Min/Max 両端部では推定が悪いことが多く留意
- RMSE 2乗平均平方根誤差
- MAE 平均絶対誤差(RMSEとは近い)
- ME 平均誤差
- AIC 赤池情報量基準 ……など

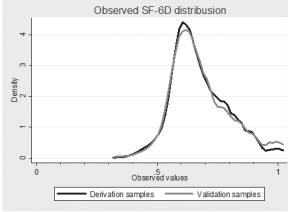
効用値の各区間毎の当てはまりの評価を実施。
外的基準(併存疾患数)を用いた場合の妥当性検証。

スライド-10

10

結果：対象者の特性

	Derivation dataset (n=1539)	Validation dataset (n=769)
年齢,歳(SD)	51.0 (16.2)	51.5 (15.4)
女性,%	49.4	47.9
SF-6D (SD)	0.68 (0.12)	0.68 (0.13)
糖尿病,%	7.6	7.0
高血圧,%	19.7	22.4
心血管疾患,%	3.6	2.9
脳血管疾患,%	2.2	1.6
併存症数 (SD)	0.72 (0.98)	0.75 (1.00)



Observed SF-6D distribution

Density

Observed values

Derivation samples Validation samples

示しています。英国の結果と比較して、ばらつきに関しても、多少点数は低いものの概ねばらついていると判断し、マッピングは可能だろうという判断に至りました。

【スライド-11】

次は初期モデルの結果からの抜粋ですが、最も当てはまりのよかった2つのモデルを提示しています。結果、dummy variable、つまりダミー変数で入れたモデルの当てはまりが最も良好でした。その中で最尤推定法に基づく線形回帰とTobitモデルの結果を示しています。ただ、少し図で分かりにくいかもしれませんが、誤差を見ると、高い推定値においては推定値との差が大きくなっており、当てはまりが悪いことがそれぞれ示されています。ここを改善する必要があるということが初期モデルからは明らかになりました。なお、さまざまな当てはまりの指標を見ますと、最大値以外は線形回帰分析のほうがよろしいので、そちらを使用することとしました。

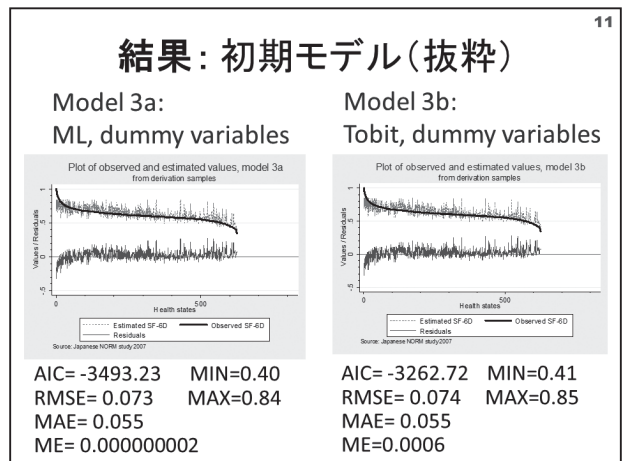
【スライド-12】

そのモデルを拡張したモデルの結果をお示しします。これは、性、年齢、さらには先ほど申しましたMOST termsを入れて、2段階で用いたものです。やはりまだ、こちらのよい健康状態のところの当てはまりは少し悪いのですが、だいぶ改善されてきて、最大値は0.90になっています。そしてまた、validation data setでも良好な当てはまりが確認され、本モデルの使用は問題ないであろうという判断に至りました。

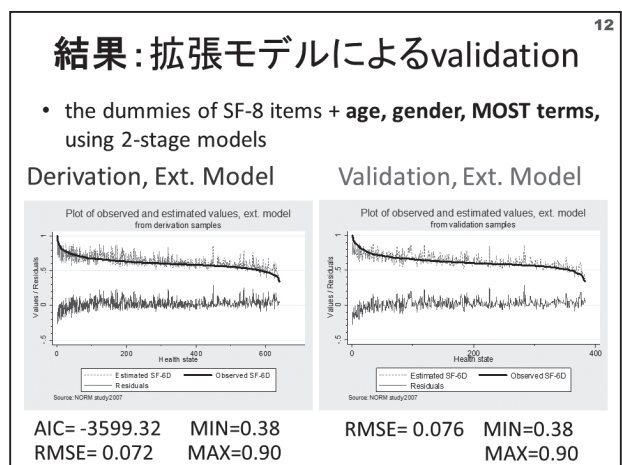
【スライド-13】

考察です。

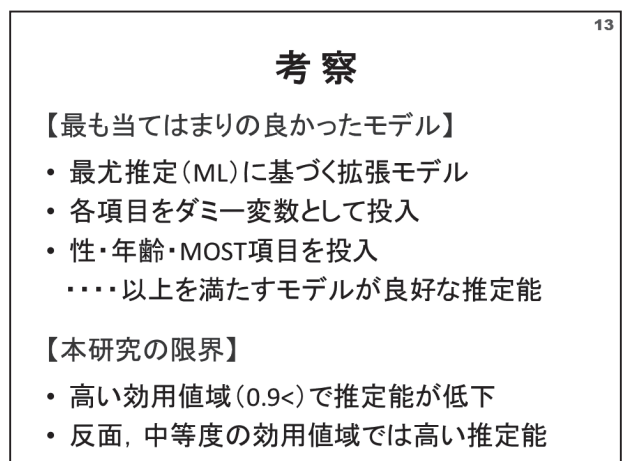
スライド-11



スライド-12



スライド-13



最尤推定の線形回帰分析に基づく拡張モデルで、各項目をダミー変数として投入し、さらには性、年齢、MOST termsを投入したモデルが良好な推定能を満たすということが明らかとなりました。

ただ、本研究の限界としては、高い効用値域、特に0.9以上の効用値域においては推定能が低下するという問題点があり、すなわち非常に健康状態の良好な方に対して本研究より得られたモデルは使いにくい点が挙げられます。反面、中等度域のものに関しては高い推定能を有していますので、病院の入院患者さん等に関しては十分使用可能であろうということが示唆されました。

【スライド-14】

結語です。

SF-8の項目から、効用値を換算可能な推定式の算出を行いました。

性、年齢などの変数投入、ならびに2段階のモデルの使用により、推定能の向上に成功しました。今後さらなるモデルの改良を予定しています。

スライド-14

結語

14

- SF-8の項目から、効用値を換算可能な推定式の算出を行った。
- 性・年齢などの変数投入ならびに2段階のモデルの使用により、推定能の向上に成功した。
- 高い効用値域での推定能の向上に向けて、今後更なるモデルの改良を予定。

質疑応答

座長： この推定結果を利用した研究は、かなり進んでいるのですか？

山本： はい、推定の論文を執筆中でして、その結果の公表され次第、恐らく本推定方法を用いた研究が拡がるものと思われます。すなわち、短い項目の質問紙からでも効用値が算出できるようになるというメリットを社会に還元できるものと思います。

座長： データベースをきちっと整理しておけばしておくほど、将来非常に面白い研究ができますね。

山本： フィードバックして、その結果をまたさらに分析するというを繰り返していくことで、どんどん精度が高まってくると思われますので、そのようなことを期待しているところでございます。

座長： ぜひ頑張ってください。

山本： ありがとうございます。