

令和 2 年 7 月 2 日現在

機関番号：14301

研究種目：若手研究(B)

研究期間：2017～2019

課題番号：17K15744

研究課題名（和文）日本の医療ビッグデータに基づく効用値・QOL推定アルゴリズムの作成と検証

研究課題名（英文）Estimating utilities and health-related QOL based on the real-world data in Japan

研究代表者

山本 洋介 (yamamoto, yosuke)

京都大学・医学研究科・准教授

研究者番号：30583190

交付決定額（研究期間全体）：（直接経費） 4,500,000円

研究成果の概要（和文）：本研究の目的は、既存の大規模データベースに含まれる項目から、効用値を推定するアルゴリズムの開発することである。具体的には、研究代表者がアクセス可能なコホート研究のデータ、ならびに一般住民集団を対象とした新規調査を実施、上記と同様の効用値推定が一般住民においても可能かを検討した。結果、3500名を対象とした新規調査で、特定健診項目、併存疾患、性・年齢を加えたモデルについて、自由度調整済み決定係数0.317、二乗平均平方根誤差0.078であった。限界はあるものの、特定健診項目から効用値を推定するモデルを開発した。高得点域ならびに低得点域の推定精度が十分ではなく、今後更なるモデルの改善が必要である。

研究成果の学術的意義や社会的意義

本研究の結果、推定精度には一定の限界はあるものの、特定健診の結果から効用値を推定することに成功した。このことにより、日本全国のすべての保険診療ならびに特定健診のデータベースである、レセプト情報・特定健診等情報データベース（NDB）のデータから効用値を推定できる可能性が示されたと言え、これまでQOLのデータがないことが限界のひとつであった、大規模医療リアルワールドデータのさらなる活用につながることで期待される。具体的には、効用値を用いた医療経済評価分析などを通じて、今後の日本の医療行政を進める上で不可欠な分析に寄与しうる点に、社会的意義があると思われた。

研究成果の概要（英文）：The aim of the study is to develop models to estimate utilities from the variables included in real-world database in Japan. We used data of existing cohort study based on health check-ups in two municipalities in Fukushima, and newly conducted an additional survey using representative samples from Japanese general population. In the new survey, 3500 participants were enrolled. We performed regression models to estimate utilities using variables below: age, gender, comorbidities, and other variables which had been collected in specific Health Checkups in Japan. In the best performing model, the adjusted R-square was 0.317 and the root mean square error was 0.078. Further study was required to accurately predict utilities at either end of observed values.

研究分野：臨床疫学，QOL研究

キーワード：効用値 QOL 医療ビッグデータ 健康関連QOL尺度

様式 C - 19、F - 19 - 1、Z - 19 (共通)

1. 研究開始当初の背景

- ・ 効用値 (utility) とは、人の主観に基づく健康状態に関して、0-1 の値で価値付けを行う健康関連 QOL 尺度である。
- ・ 昨今、大規模データベースを用いた医療経済評価研究の重要性が認識されつつあるが、その研究実施には、対象者の健康状態を効用値に基づき評価することが不可欠である。
- ・ しかしながら、日本の大規模データベースは、概ねレセプトデータ・DPQ (Diagnosis Procedure Combination; 診断群分類) データ・特定健診データに基づくため、効用値を直接推定できない。
- ・ そこで本研究では、既存の大規模データベースに含まれる項目から、効用値を推定するアルゴリズムの開発を行う。本研究により、日本発の医療データベース研究に QOL 情報を付加することが可能となり、医療経済評価研究の質向上に大きく寄与することが期待される。

2. 研究の目的

1) 大規模データベースに含まれる項目から効用値を推定するアルゴリズムの開発

本研究では、研究代表者がアクセス可能なコホート研究のデータを用いて、大規模データベースに一般的に含まれる項目から、一意の効用値を推定するモデルを開発することを目的とした。

2) 日本の代表サンプルにおける効用値推定のアルゴリズムの開発

また、一般住民パネルから、日本の性・年齢・地域の構成比に即したサンプリングを行ったうえで、一般住民集団を対象として、効用値の direct な測定を含む新規調査を実施、上記と同様に汎用性の高い項目から効用値推定が一般住民全体でも可能かを検討することを目的とした。

3. 研究の方法

1) 既存コホートに含まれる項目から効用値を推定するアルゴリズムの開発

研究者がアクセス可能なコホートデータの中の特定健診に係る項目を中心に、効用値を推定するモデルを複数作成し、比較検討を行った。

2) 新規調査に基づくアルゴリズムの開発

- ・ さらに、日本の代表性を担保された一般住民集団を対象とした新規質問票調査を実施した。その調査には、特定健診で測定されている問診項目も含めることにより、上記 1) の既存コホートデータと組み合わせることで、derivation ならびに validation を実施した。
- ・ また、リファレンスとなる効用値の推定に用いるインデックス型尺度として、SF-6D に加え、EQ-5D-5L も用いた。

3) 皮膚疾患を対象とした直接効用値の利活用性の検討

- ・ インデックス型尺度の性質上、十分な効用値の測定が困難であると考えられる健康状態のひとつに、皮膚疾患が挙げられる。そこで、Composite Time Trade-off (TTO) 法にのっとり、self-TTO (患者の実際の健康状態の TTO 法に基づく価値づけ) による効用値推定を試みた。具体的には皮膚疾患を有する患者を対象として、上記のインデックス尺度によって得られた効用値と、自身の健康状態に関する self-TTO とを比較検討を行った。

(実際のモデル作成の手順)

- ・ 既存コホートデータを用いて、特定健診に含まれる多数の項目から、臨床的変数選択を第一に、必要に応じてステップワイズ法 (変数減少法) などにより説明変数を抽出した後、コホートの質問項目に含まれるインデックス型尺度により得られた効用値を従属変数とした回帰モデルを作成した。
- ・ 引き続き、日本の一般住民集団に基づく新規調査の結果を用いて、上記と同様の手順で複数のモデルを作成した。
- ・ その際の従属変数は SF-6d・EQ-5D-5L から得られた効用値とし、説明変数としては、特定健診の問診項目に加えて、性・年齢・併存疾患など、ほぼすべての大規模データベースでも活用可能な因子も加えて推定を行った。
- ・ なお、皮膚疾患を有する対象者においては、EQ-5D などの推定の際に用いる cTTO に準拠し、直接自身の皮膚の状態を効用値として答えてもらい、皮膚の疾患特異的尺度である Dermatology Quality of Life Index (DLQI) による QOL の総スコアに対する推定性能について、インデックス型尺度と比較した。
- ・ 上記のモデル推定には、検証にはベイズ情報量基準 (BIC) ならびに自由度調整済み決定係数の他、root mean square error (RMSE) や MAE (mean absolute error) などを用いて、推定効用値と実際の効用値の間の誤差の大きさ、系統的誤差の有無などを検証し、推定効用値を求めるためのモデルを選択した。
- ・ 上記の手順で十分な推定能が得られない場合において、機械学習モデル (ランダムフォレストモデル、ならびにニューラルネットワークモデル) を用いた推定を試行、推定能の改善が得られるかについての検討を追加した。

4. 研究成果

1) 既存コホートに含まれる項目から効用値を推定するアルゴリズムの開発について

・研究者がアクセス可能な福島県山間部の既存コホート(2008年度調査)を用いた。この研究では、40歳以上の当該地域に住む特定健診を受検した男女3218名(平均年齢66.5歳(標準偏差10.4)、男性40.0%)が対象となった。

・まず、日本語版SF-36からSF-6dに基づく効用値を算出した上で、効用値を従属変数、自覚症状8項目と問診項目10項目を説明変数としたベースとなる重回帰モデルを作成したところ、自由度調整済み決定係数0.097、RMSE=0.120であった。

・さらに、併存疾患ならびに性・年齢を加えた重回帰モデルにおいても、自由度調整済み決定係数0.130、RMSE=0.119と若干の改善を認めたものの、推定式の作成に耐えうるものとはいえなかった。

・そのため、機械学習(ニューラルネットワーク・ランダムフォレスト)の手法を用いて推定を試みた。2:1のスプリットサンプルによるテストデータを用いて推定能を確認したところ、各モデルのRMSEはいずれも0.116前後と推定能の十分な改善は認められなかった。

2) 新規調査に基づくアルゴリズムの開発について

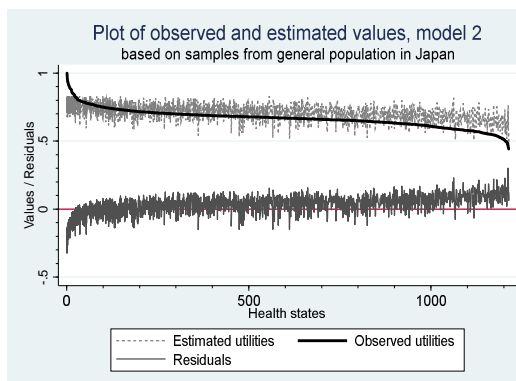
・既存コホートでは十分な推定能が得られなかったことを踏まえ、日本の性・年齢・地域区別の人口構成比に割り当てられた一般住民集団3500名(平均年齢48.8歳(標準偏差16.7)、男性49.5%)に対して新たに調査を実施した。

・まず、日本語版SF-36からSF-6dに基づく効用値を算出した上で、効用値を従属変数、自覚症状9項目と問診項目10項目を説明変数とした重回帰モデルを作成したところ、自由度調整済み決定係数0.291、RMSE=0.080であった。

・さらに、併存疾患ならびに性・年齢を加えた重回帰モデルにおいては、自由度調整済み決定係数0.305、RMSE=0.079と改善を認めたため、このモデルを元に複数のモデルを作成してそれらの性能を比較検討した。

・複数モデルにおける検討の結果、restricted cubic splineを用いたモデルにおいて、自由度調整済み決定係数0.317、RMSE=0.078と若干の改善を認めた(右図)。ただし、効用値の高い領域ならびに低い領域については良好な推定精度を担保できておらず、今後の課題となった。

・また、リファレンスとなる効用値の推定に用いるインデックス型尺度として、EQ-5D-5Lに基づく効用値を従属変数としたモデル作成を行った。具体的には、EQ-5D-5Lの値を従属変数とし、自覚症状9項目、問診項目10項目、併存疾患ならびに性・年齢を説明変数とした重回帰モデルにおいては、自由度調整済み決定係数0.301、RMSE=0.120と、先の推定と比較してその精度は若干低い結果となった。



3) 皮膚疾患を対象とした直接効用値の利活用性の検討について

・新たに実施した調査研究のうち、310名は皮膚疾患を有するとの回答が得られた。self-TT0は平均0.698(標準偏差0.45)であり、SF-6dならびにEQ-5Dに基づく効用値より低く、標準偏差は大きい傾向を示した。なお、DLQIの合計スコア得点(範囲0-30点)の平均は、3.8点であった。

・DLQIのスコア得点を従属変数とし、それぞれの効用値を説明変数とした各モデルの比較を行ったところ、自由度調整済み決定係数はそれぞれself-TT0で0.131、SF-6dで0.134、EQ-5Dで0.118であった。本研究の結果より、皮膚疾患に絞ったTT0を測定が、皮膚疾患のQOLの状態の反映という観点において、他の既存の尺度と比較しても優れているとはいえなかった。self-TT0の測定については十分に説明を加えたものの難しい点が多く、今後さらなる測定法の検証が必要である。

<参考文献>

・Shiroiwa T, Ikeda S, Noto S, Igarashi A, Fukuda T, Saito S, Shimozuma K. Comparison of Value Set Based on DCE and/or TTO Data: Scoring for EQ-5D-5L Health States in Japan. *Value Health*. 2016;19: 648-54.

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 0件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 山本洋介	4. 巻 71
2. 論文標題 患者立脚型調査を用いた研究	5. 発行年 2020年
3. 雑誌名 臨床整形外科	6. 最初と最後の頁 735-738
掲載論文のDOI（デジタルオブジェクト識別子） https://doi.org/10.15106/j_seikei71_735	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計1件（うち招待講演 0件 / うち国際学会 0件）

1. 発表者名 山本洋介, John Brazier, Clara Mukuria, 大西良浩, 福原俊一
2. 発表標題 日本語版SF-8に基づく効用値算出アルゴリズムの開発
3. 学会等名 日本臨床疫学会 第1回年次学術大会
4. 発表年 2017年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----