

脳卒中の障害状態についての効用値の評価

評点尺度法と時間得失法による検討

ノト シンイチ ヤナギ ヒサコ トムラ シゲオ
能登 真一* 柳 久子^{2*} 戸村 成男^{2*}

目的 費用効用分析でその結果として用いられる効用は、一般に質を調整した生存年 (quality-adjusted life years: QALYs) で表される。QALYsはある健康状態について、完全に健康な状態を1、死亡を0とする尺度で効用値 (utility) を数量化し、それに生存年数を乗じて求められる。リハビリテーションで多く関わる脳卒中は様々な障害状態を残存させるため、その効用値は障害状態ごとに求められなければならない。今回、脳卒中の障害状態のスケールとして最も広く用いられているものの一つである Rankin scale による障害状態ごとの効用値を求め、その評価に関わる人口学的特徴の影響や測定方法の関係の分析を目的とした。

方法 効用値の評価は、評点尺度法と時間得失法の2方法とし、それぞれ質問紙法にて測定した。評点尺度法は一方の端に完全に健康な状態を、もう一方の端に死を置いた線分上に複数の健康状態を位置付ける方法で、時間得失法は悪い健康状態に対して良い健康状態を得る際にどのくらいの時間をあきらめてもよいかを問う方法である。対象は、大学生、リハビリテーションスタッフ、脳卒中患者の介護者、一般企業会社員の計460人である。統計的手法はノンパラメトリック法を用いた。

成績 Rankin scale I は評点尺度法で0.89、時間得失法で0.83、以下、Rankin scale II はそれぞれ0.72と0.67、Rankin scale III は0.56と0.45、Rankin scale IV は0.36と0.24、Rankin scale V (寝たきりの状態) は0.18と0.09となった。評点尺度法と時間得失法の相関は、0.176~0.412の範囲となった。時間得失法による Rankin scale I、III~V で集団間に差を認めたが、年齢、性別、学歴、婚姻関係、健康状態などの人口学的特徴によって効用値に有意な差を認めたものは僅かであった。

結論 脳卒中の障害状態についての効用値は、障害状態により大きく異なった。効用値の測定方法については、評点尺度法と時間得失法の相関はそれほど強いものではないことが実証された。また、健康状態の選好には人口学的特徴の影響が少ないことが明らかとなった。

Key words : 脳卒中、効用値、評点尺度法、時間得失法、QALYs、リハビリテーション

I 緒 言

臨床経済学は、医療サービスという資源を配分する際に効率性の観点から評価する学問であるが、その手法には費用便益分析、費用効果分析、費用効用分析の3つの方法がある。中でも費用効用分析は、費用便益分析の余命を金銭換算したり、費用効果分析の他の治療プログラムとの比較

ができないといった問題点を克服した評価法としてその有用性が訴えられ^{1,2)}、近年ますますその利用が増加している^{3,4)}。

費用効用分析で用いられる結果の尺度は、『質を調整した生存年』(Quality-Adjusted Life Years: QALYs) が用いられる。QALYsは効用値 (utility) に生存年を掛け合わせることで求められるが、その意味するところは質を加味した生存年であり、一つの指標で生命の質と量の両面をとらえている。この指標を用いれば、あらゆる医療サービス間での対費用効果の比較が可能になる。実際には、各治療方法についてそれらの1QALY延長に要する費用の増分を比較することで、経済的効

* 新潟医療福祉大学医療技術学部作業療法学科

^{2*} 筑波大学社会医学系

連絡先：〒950-3198 新潟市島見町1398

新潟医療福祉大学医療技術学部作業療法学科

能登真一

果の評価や政策決定に用いられている。そこで用いられる効用値は生活の質を単一の指標として評価するもので、ある健康状態の価値を完全に健康な状態を「1」、死を「0」とした尺度の中で、患者あるいは一般対象者の価値判断により測定するものである⁵⁾。

さまざまな健康状態についての効用値の測定方法については、時間得失法や基準的賭け法、評点尺度法などで標準化されてきているものの、各測定方法間で必ずしも統一された合意は得られていない⁶⁾。一方、これらの手法は欧米で開発され、発展してきたものでその方面では研究も盛んであるが、本邦における研究は久繁^{7,8)}、田村ら⁹⁾を除いてほとんどない。臨床経済学は様々な立場から実施されているが、本来は公共政策の意志決定と関連するものであるため、総合的な社会的立場で実施されるのが望ましく、適切な効用値もその国々において測定されたものでなければならぬはずである。今後、本邦においても臨床経済学の立場から治療プログラムが評価され、その結果が政策決定に影響を与えていくと考えられるが、その前段階として効用値を実証的に測定し、その妥当性を検討することは重要と考えられる。

二木^{10~12)}はリハビリテーション医療の効果に関して、ADLに加えてQOLを多面的・科学的に評価する必要があることを指摘し、臨床経済学、とくに費用効用分析の有効性について言及している。本邦でも近年になって、臨床経済学的手法による検討が行われているが、それらは費用便益分析^{13,14)}や費用効果分析^{15,16)}などで、費用効用分析による検討はまだなされていない。川久保¹⁷⁾も心臓リハビリテーションについて、費用効用分析の必要性を述べている。

リハビリテーション医療の中で最も多く関わるのは脳卒中であるが、脳卒中はさまざまな障害状態を呈するため、効用値もその障害状態ごとに設定されなければならない。脳卒中の障害状態の効用値については、これまで海外においてさまざまな研究がなされている。測定手法は時間得失法^{18~23)}、基準的賭け法^{24~26)}、評点尺度法^{27~29)}により直接測定されたものが多く、一部はHealth Utility Index^{30,31)}やEuroQol³²⁾といったカテゴリー分類によって測定されたものもある。対象者は健常人や脳卒中リスク患者、患者本人に至るま

でさまざまである。脳卒中の障害状態の設定については、minor stroke, major stroke という大まかな分類になっているものが多い。近年、それらの研究がレビューとしてまとめられ、Tengs³³⁾によれば minor stroke 0.45~0.92, moderate stroke 0.12~0.81, major stroke -0.02~0.71の範囲であり、Postら³⁴⁾によれば脳卒中リスク患者による評価で minor stroke が0.55, major stroke が0.26と報告された。

このような現状からリハビリテーション医療に用いることのできる脳卒中の障害状態についての効用値を考えると、ある一定の障害スケールに基づいた効用値の評価が求められる。これはリハビリテーション医療で、機能障害 (impairment)、能力障害 (disability)、社会的不利 (handicap) といった障害の改善が一方で重要とされているため、その改善度に合わせた効用値が有効と考えられる。その脳卒中の障害状態のスケールとして最も広く用いられているものの一つに Rankin scale³⁵⁾があり、5つのグレードごとに障害状態が分類されている。この Rankin scale に従い、そのグレードごとに効用値が設定されれば、そのグレードの改善から患者ごとの効用値の差が求められ、ひいてはそれが QOL の改善を表す費用効用分析の結果の尺度として用いられることになると考えられる。

このように、本研究は、脳卒中リハビリテーション医療の費用効用分析を行うための前段階として、Rankin scale に基づいた脳卒中の障害状態ごとの効用値を評点尺度法と時間得失法を用いて評価し、その妥当性とその評価に係る因子について検討することを目的として行った。

II 研究方法

調査は2000年6月から9月にかけて、18歳から70歳までの健康な成人を対象に行った。今回の調査はサンプル数をできるだけ短期間に多く集めることと脳卒中の障害状態の理解が得やすいことを主な条件に対象者を選定した。また、すでに一定の障害状態にある患者については、倫理上の理由から今回の研究の対象から除外した。この結果、対象は筑波大学医学部の学生 (以下、学生群) 171人、脳卒中リハビリテーションを施行している病院に勤める理学療法士・作業療法士・言語聴

覚士のリハビリテーションスタッフ（以下、リハスタッフ群）154人、茨城県のT病院に入院もしくは外来でリハビリテーションを施行されている脳卒中患者の家族介護者（以下、介護者群）55人、さらに某生命保険会社に勤める会社員（以下、会社員群）160人、合計540人となった。リハスタッフ群の対象者は福岡県、神奈川県、静岡県、長野県、福岡県の脳卒中リハビリテーションを専門に行っている病院からボランティアを募った。同様に、会社員群の対象者は福岡県、東京都、長野県、島根県の営業所に勤める社員からボランティアを募った。

人口学的特徴は、性別、年齢、生活状況（一人住まい、家族住まい）、結婚状況（未婚、既婚）、現在の健康状態（とても良好、ほぼ良好、やや不良、不良）、最終学歴（中学校、高校、短大・専門学校、大学）、年間収入（300万円以下、301～500万円、501～700万円、701万円以上）、高齢者との同居の経験の有無、障害者との同居の経験の有無についてたずねた。年齢はさらに18～24歳、25～34歳、35～44歳、45～70歳の4群に分けた。

調査方法についても、できるだけ短期間に多くのサンプルを得るために自記式の質問紙法を用い、無記名で行った。質問紙の配布と回収は学生群と介護者群には手渡しにて、リハスタッフ群および会社員群については郵送にて行った。また、それぞれの対象者に対し口頭、もしくは文書で研究の目的を説明し、同意を得た上で調査を行った。

脳卒中の障害状態は、Rankin scaleに基づき5段階の障害レベルとした。すなわち、Rankin scale Iは麻痺なし、Rankin scale IIは不完全な片麻痺で介護者が不要なレベル、Rankin scale IIIは不完全な片麻痺で杖歩行レベル、Rankin scale IVは完全な片麻痺で車椅子レベル、Rankin scale Vは完全な片麻痺で寝たきり状態である。これらを対象者に理解しやすいよう質問紙の中にシナリオとして提示した（表1）。対象者はこれらを読んで効用値の判断を行った。

効用値の測定方法には、評点尺度法と時間得失法を用いた。評点尺度法は視覚的な尺度を用いて評価する方法である^{36,37}。尺度には両端がはっきりと定められた線分を用いる。最も好ましい健康状態をこの線分の一方の端に置き、最悪の健康状態をもう一方の端に置く。対象者はこの両端の間に評価の対象となる健康状態を好みの順に並べ、線上の間隔と健康状態の価値が対応するように位置づける。また、この方法では、しばしば対象者の理解を助けるために感情温度計と呼ばれる道具が用いられている³⁸。今回は図1に示すような感情温度計を作成し、そこに5つの障害状態をその健康状態の好ましさに応じて直接矢印で図示するよう求めた。この測定方法による効用は矢印で示された値を100で除ことで求められる。

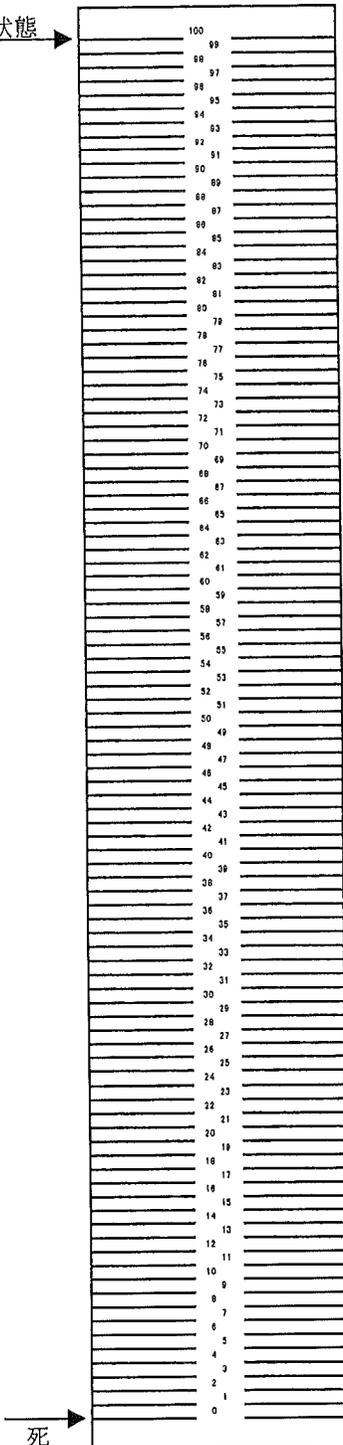
時間得失法は、保健医療の分野での使用を目的として、Torranceら³⁹によって考案された測定方法である。それは、ある障害状態に変えて、ま

表1 Rankin scaleに基づく障害状態のシナリオ

Rankin scale (分類尺度)	
Grade I	左右どちらかの手足にごく軽い麻痺があったが、ほとんど改善し、日常生活や仕事もこれまで通り問題なく行うことができる。
Grade II	左右どちらかの手足の不完全な麻痺のため、日常生活や仕事でこれまでより時間がかかるなど支援があるが、介助者は必要ない。
Grade III	左右どちらかの手足の不完全な麻痺のため、歩く際に杖や装具が必要であるが介助なしに歩ける。しかし、入浴や着替えの際には介助が必要である。話す言葉は不明瞭だが日常のコミュニケーションに支障はない。
Grade IV	左右どちらかの手足の完全な麻痺のため、介助なしには歩けず、日中のほとんどを車椅子かベッドの上で過ごす。食事は自分でとることができるが、排泄の際にも介助が必要である。言葉を話したり、文字を書いたりすることにも支障がある。
Grade V	左右どちらかの麻痺と意識障害のため、完全にベッドで寝たきりの状態で、食事や排泄の際にも介助が必要である。言葉を話したり、文字を書くことも難しく、他人とコミュニケーションをとることがほとんどできない。

図1 感情温度計

完全に健康な状態 →



まったく健康な状態を得るためにどれくらいの余命をあきらめても良いかを問う。つまり、ある障害状態で何年か生存できることと、完全に健康な状態でそれよりも少ない何年かを選択していくものである。今回は、図2に示す例を提示した上で、「あなたが60歳で脳卒中を発症し、ある障害状態になったと仮定します。この障害状態xで20年間生きることと、寿命を短くして何年間か完全に健康な状態を得て生きることのどちらかを選択できることとします。」と質問し、障害状態xで20年生きることと完全に健康な状態で生きることの2つの選択肢を用意した上で、完全に健康な状態で生きられる年数を徐々に減少させていき、それぞれどちらかを選択してもらうこととした。すなわち、この方法では、ある障害状態で20年生きることと完全に健康な状態を得た上で20年より少ない年数を生き方のどちらかを選択するその分岐点を調べ、それをその障害状態の健康価値とする。例の障害状態Xの場合、分岐点が9年であるから効用値は9年/20年で0.45となる。この測定方法では、対象者の理解を助けるために視覚的補助(得失板)^{37,40)}を用いることが試みられているが、今回の研究では一部の対象者に郵送法を用いたため得失板は使用しなかった。

また、ある障害状態で生存できる期間を20年としたことについて、先行研究ではその期間の設定が1年⁴¹⁾、5年⁴²⁾、10年⁴²⁾、20年^{9,43)}とさまざまであるが、平成11年簡易生命表⁴⁴⁾によれば、60歳の日本人の平均余命は男性で20.9歳、女性で26.3歳であり、今回のシナリオを理解する上では妥当な期間と考えた。

この評価方法の信頼性については、茨城県下の病院に勤めるリハビリテーションスタッフ23人を対象にした先行研究において、再テスト法によって検討し信頼性が確認された⁴⁵⁾。

統計学的分析はすべてノンパラメトリック法を用いた。これは今回のデータの分布がRankin scaleのグレード毎に異なり、その一部が正規分布に従っていなかったためである。これまでの効用値測定に関する研究では、比較的例数が多い場合は正規分布を前提としても問題ないという意見⁷⁾があり、かつ、ノンパラメトリック法はそもそも数値を検定の対象にしない点で検出力が劣るが、今回はパラメトリックな手法を用いることに

図2 時間得失法の質問と回答の例

【例】あなたが60歳で脳卒中を発症し、ある障害状態Xで20年間生きることより、完全に健康な状態で10年生きることの方がよいと考えた場合（完全に健康な状態でも8年しか生きられないのでは少ないと考えた場合）、以下のように印を付けてください。

それぞれどちらを選択しますか

障害状態Xで20年生きる	それとも	完全に健康な状態で18年生きる
障害状態Xで20年生きる	それとも	完全に健康な状態で16年生きる
障害状態Xで20年生きる	それとも	完全に健康な状態で14年生きる
障害状態Xで20年生きる	それとも	完全に健康な状態で12年生きる
障害状態Xで20年生きる	それとも	完全に健康な状態で10年生きる
障害状態Xで20年生きる	それとも	完全に健康な状態で8年生きる
障害状態Xで20年生きる	それとも	完全に健康な状態で6年生きる
障害状態Xで20年生きる	それとも	完全に健康な状態で4年生きる
障害状態Xで20年生きる	それとも	完全に健康な状態で2年生きる
障害状態Xで20年生きる	それとも	完全に健康な状態で0年生きる（死に等しい）

よる結論の信頼性低下を防ぐ目的を重視した。評点尺度法と時間得失法の相関は Spearman の順位相関係数を、2方法の差の比較は Wilcoxon の順位テストを用いた。Rankin scale のグレード毎の比較には、Friedman の順位検定を、対象者の人口学的特徴によって群分けした効用値の比較には Kruskal-Wallis 検定および Mann-Whitney の U 検定を用いた。これら統計処理は、すべて Stat-View version 5.0 (Abacus Concept, Inc. Berkeley, CA, USA) で行った。

III 結 果

質問の有効回答数は学生群169人（回収率98.8%）、リハスタッフ群134人（87.0%）、介護者群52人（94.5%）、会社員群105人（65.6%）、総数で460人であり、対象者全体の540人に対する回収率は85.2%であった。

表2に対象者の人口学的特徴を示す。全体の男女比は男性197人、女性263人であった。学生群、リハスタッフ群、介護者群、会社員群の平均年齢は、それぞれ19.2歳、29.5歳、40.9歳、42.2歳で全体では29.9歳であった。

Rankin scale ごとの効用値を表3に示す。評点尺度法による Rankin scale I は平均値が0.89、II は0.72、III は0.56、IV は0.36、V は0.18となった。Rankin scale のグレードごとに有意な差を認めた ($P < 0.0001$)。対象者が Rankin scale のすべてのグレードで1.00と答えた者はリハスタッフ群の1人、介護者群の3人、会社員群の4人であった。

また、Rankin scale のすべてのランクについて0と答えた者は会社員群の1人のみであった。それ以外の対象者はすべて Rankin scale のグレードが上がるごとに効用値を低く回答した。

時間得失法による平均値は Rankin scale I から順に0.83, 0.67, 0.45, 0.24, 0.09となった。評点尺度法と同様に時間得失法でも Rankin scale のグレードごとに有意な差を認めた ($P < 0.0001$)。対象者が Rankin scale のすべてのグレードで1.00と答えた者はリハスタッフ群の4人、会社員群の1人であり、Rankin scale のすべてのグレードについて0と答えた者は会社員群の4人のみであった。それ以外の対象者は評点尺度法と同様、すべて Rankin scale のグレードが上がるごとに効用値を低く回答した。

Rankin scale ごとの評点尺度法と時間得失法の2方法における相関関係については、Rankin scale I で r 値が0.165 ($P = 0.0004$)、Rankin scale II で0.203 ($P < 0.0001$)、Rankin scale III で0.215 ($P < 0.0001$)、Rankin scale IV で0.220 ($P < 0.0001$)、Rankin scale V では0.419 ($P < 0.0001$) となった。

次に、人口学的特徴による比較の結果を以下に示す。まず、集団による効用値の比較の結果を表4に示す。評点尺度法では、Rankin scale I からIVのグレードで集団による違いを認めなかったが、Rankin scale V では集団による有意な違いを認めた ($P = 0.0015$)。しかし、平均値でみると学生群0.19、リハスタッフ群0.18、介護者群0.17、会社員群0.18となり、その差は小さいものであった。

表2 対象者の人口学的特長

	学 生 (n=169)	リハスタッフ (n=134)	介護者 (n=52)	会社員 (n=105)	合 計 (n=460)
性別					
男性	120	34	21	22	197
女性	49	100	31	83	263
年齢					
平均年齢 (歳)	19.2	29.5	40.9	42.2	32.9
18~24歳	169	38	4	6	217
25~34歳	0	67	15	23	105
35~44歳	0	23	9	31	63
45~60歳	0	6	24	45	75
生活状況					
一人住まい	164	60	0	4	228
家族住まい	5	74	52	101	232
結婚状況					
未婚	169	88	18	15	286
既婚	0	46	34	90	174
健康状態					
とても良好	46	32	14	20	112
ほぼ良好	101	90	35	65	291
やや不良	21	12	3	20	56
不良	1	0	0	0	1
最終学歴					
中学校	0	0	6	1	7
高校	168	0	25	68	261
短大・専門学校	1	97	10	18	126
大学	0	37	11	18	66
年間収入					
300万円以下	169	56	25	40	289
301~500万円	0	63	11	32	106
501~700万円	0	14	14	11	39
701万円以上	0	1	2	22	26
高齢者との同居の経験					
あり	77	69	26	41	213
なし	92	65	26	64	247
障害者との同居の経験					
あり	11	16	12	15	54
なし	158	118	40	90	406

時間得失法では、Rankin scale IIで集団による違いを認めなかったが、Rankin scale I ($P=0.0035$), Rankin scale III ($P=0.0046$), IV ($P=0.0061$), V ($P=0.0083$) で有意な違いを認めた。時間得失法では、有意な違いを認めなかった Rankin scale II を含めて、Rankin scale I から Rankin scale III までのグレードでは集団の中で介護者群が他の群に比べて効用値を高く回答し、会社員群が

最も低く回答した。Rankin scale IV と Rankin scale V では効用値を最も高く回答したのはリハスタッフ群であった。

次に、年齢による効用値の比較を表5に示す。ここでは、年齢で4つの群に分けた。すなわち、18~24歳、25~34歳、35~44歳、45~68歳で人数はそれぞれ217人、105人、63人、75人であった。評点尺度法では、Rankin scale I から IV で年齢に

表3 効用値の測定結果と2測定方法間の相関

Rankin scale	評点尺度法*	時間得失法*	r† (P値)
I	0.89±0.11/ 0.90	0.83±0.23/ 1.00	0.165(0.0004)
II	0.72±0.15/ 0.75	0.67±0.27/ 0.65	0.203(<0.0001)
III	0.56±0.17/ 0.55	0.45±0.24/ 0.45	0.215(<0.0001)
IV	0.36±0.19/ 0.33	0.24±0.21/ 0.25	0.220(<0.0001)
V	0.18±0.18/ 0.11	0.09±0.17/ 0.05	0.419(<0.0001)

Mean±SD/Median

* Friedman's rank test (P<0.0001), † Spearman's rank correlation test

表4 集団による効用値の比較*

Rankin scale	学生 (n=169)	リハスタッフ (n=134)	介護者 (n=52)	会社員 (n=105)	P値
(評点尺度法)					
I	0.89	0.89	0.89	0.89	ns
II	0.73	0.72	0.72	0.73	ns
III	0.56	0.56	0.53	0.55	ns
IV	0.37	0.35	0.34	0.35	ns
V	0.19	0.18	0.17	0.18	0.0015
(時間得失法)					
I	0.84	0.84	0.90	0.75	0.0035
II	0.69	0.69	0.72	0.60	ns
III	0.46	0.47	0.51	0.39	0.0046
IV	0.25	0.26	0.26	0.20	0.0061
V	0.08	0.12	0.09	0.08	0.0083

* Kruskal-Wallis one-way ANOVA

表5 年齢による効用値の比較*

Rankin scale	18~24歳 (n=216)	25~34歳 (n=105)	35~44歳 (n=63)	45~60歳 (n=75)	P値
(評点尺度法)					
I	0.90	0.90	0.89	0.88	ns
II	0.73	0.72	0.75	0.71	ns
III	0.58	0.55	0.57	0.54	ns
IV	0.36	0.34	0.33	0.36	ns
V	0.19	0.17	0.16	0.22	0.0336
(時間得失法)					
I	0.84	0.83	0.81	0.80	ns
II	0.68	0.68	0.70	0.61	ns
III	0.47	0.45	0.44	0.44	ns
IV	0.25	0.24	0.21	0.24	ns
V	0.08	0.11	0.07	0.12	ns

* Kruskal-Wallis one-way ANOVA

よる違いは認められなかったが、Rankin scale Vのみ有意に年齢による違いを認めた(P=0.0336)。時間得失法では、Rankin scale IからVのすべてのグレードにおいて、年齢による違いを認めなかった。

性別による比較では、評点尺度法のRankin scale I~Vのすべてのグレードで男女に有意な効用値の差はなかった。しかし、時間得失法のRankin scale Iは男性の平均値が0.86、中央値が1.00で女性はそれぞれ0.80と0.85、Rankin scale IIは男性0.70と0.75、女性0.65と0.65となり、両グレードで有意に男性の方が高くなった(P<0.05)。

また、生活状況については、一人住まいと家族住まいの両方でRankin scale IからVのすべてのグレードについて、評点尺度法、時間得失法のどちらの測定方法によってもその効用値に有意な差を認めなかった。結婚状況についても同様に、両測定方法、いずれのRankin scaleでも未婚者、既婚者の別で効用値に有意な差を認めなかった。現在の健康状態での比較では、健康状態が不良と答えた者は1人のみであったためこの効用値は統計的解析から除外したが、両測定方法、すべてのRankin scaleのグレードで有意な差を認めなかった。最終学歴と年間収入についても、同様に両測定方法、各Rankin scaleのグレードでも有意な差は認められなかった。また、高齢者との同居の経験の有無、障害者との同居の経験の有無についても、両測定方法、各Rankin scaleのいずれにおいても、両者に有意な差を認めなかった。

IV 考 察

効用値は人々の健康状態に対する選好を表すものである。よって、その健康状態あるいは障害状態が違えば効用値も違ったものとなる。今回のRankin scaleを用いた効用値の評価において、評点尺度法、時間得失法の双方でそのグレードごとに効用値に差を認めたことは、人々の健康状態の選好が移動自立の有無やADLにおける介助の有無によって大きく左右され、それらがQOL決定に影響していることを裏付けたものと考えられる。

脳卒中の障害状態についての効用値の報告は、Tengsら³³⁾、Postら³⁴⁾によりレビューが行われている。Tengsらは67の文献から161の脳卒中に関するQOLの評価を集め、minor strokeで0.45~

0.92, moderate stroke で0.12~0.81, major stroke で-0.02~0.71の範囲になったと報告した。測定方法ごとでは, major stroke についての中央値が基準的賭け法で0.51, 時間得失法で0.32, 評点尺度法で0.23と違いを認めている。また, Post らは対象者の属性の記載がある23の文献をレビューし, minor stroke (Rankin scale II~III) について0.50~0.70, major stroke (Rankin scale IV~V) について0~0.30という効用値を報告し, さらに, 時間得失法を用いた脳卒中リスク患者による評価では minor stroke 0.55, major stroke 0.26と報告した。これらの報告をもとに今回測定した効用値を比較検討する。Rankin scale との対照がある Post らとの比較では, 今回の調査の Rankin scale II~III が評点尺度法により0.56~0.72, 時間得失法により0.45~0.67となりいずれも近似した結果となった。Rankin scale IV~V についても, それぞれ0.18~0.36, 0.09~0.24となりほぼ近い値となった。これらのことは, 地域や人種, 対象者が違って, ある健康状態についての効用値に大きな違いは認められないことを示唆するものである。

人口学的特徴による影響については, Drummond ら³⁷⁾が健康状態の選好については個々の人間の間で大きく違っていて, しかもその差は, 通常的人口学的特徴(例えば, 性, 年齢, 社会経済的地位, 宗教, 病気, 職業など)からは説明できないと述べている。また, Froberg と Kane⁴⁶⁾がいくつかの効用値算出に関する文献をレビューした中で人口学的影響については無いかあっても僅かであると述べている。今回の研究では人口学的特徴による影響はほとんど受けなかった。これは, 先の報告の一部を支持する結果ととらえることができる。

しかし, 対象者の属性により効用値に差を認め, とくに時間得失法で測定した効用値に群間差の影響が強く出て, 介護者群とリハスタッフ群が効用値を高く回答した。対象者の属性による影響に関しては, 先の文献レビューを行った Post らによると一般的に脳卒中患者が脳卒中リスク患者や健常人に比べ効用値を高く評価するとされている。また, Mathias ら³⁰⁾は脳卒中患者とその介護者を対象にした効用値の測定で, 両者に差を認めず, 介護者が患者本人に代わって効用値を回答できると述べている。今回の結果もこれらの報告を

支持するものと考えられる。

しかし, 将来的に政策決定の段階で誰の効用値を基準にするかといった問題については, Boyd ら⁴⁷⁾による患者の効用値を用いるべきという意見や Dolan ら⁴⁸⁾による一般大衆の効用値を測定すべきなどとさまざまな議論がある。Post らは, ベッドサイドにおける意志決定に用いられる選好では脳卒中リスクのある患者の効用値が最も適当であると見なしている。一方, Gold ら⁶⁾は政策決定のための費用効果分析で用いられる効用値は一般的市民からのサンプルによる選好が最適であるとしている。今回, 社会的な立場に近いと考えられる会社員による効用値が低くなったことは, 対象者の選定の際に効用値が用いられる場面の設定に関する注意が必要であることをうかがわせる。

測定方法について, 今回の研究では評点尺度法と時間得失法を用いた。これは種々ある効用値の測定方法の中で, 対象者にとって最も物理的負担が少ない測定方法として評点尺度法を, また, 欧米で最も広く用いられている測定方法として時間得失法をそれぞれ採用した。これら効用値の測定方法の妥当性については, これまでさまざまな報告があるが, 相関や回帰によってその検討を行っているものが多い⁴⁹⁾。

評点尺度法と時間得失法の2測定方法間の相関関係については, まず Torrance⁵⁰⁾が一般人に慢性の健康状態について評価した中で, 評点尺度法と時間得失法の相関は0.40と弱いものであると報告した。また, Churchill ら⁵¹⁾は慢性の腎疾患患者について調査した中で, その相関を0.22と報告した。さらに, 脳卒中のリスクの高い患者群を対象に3つの脳卒中後の障害状態について効用値を調査した Samsa ら²⁰⁾の研究では評点尺度法と時間得失法の相関は0.41であった。今回の結果は, 2測定方法間の相関が0.165~0.419となり, 先の研究とほぼ近い値となった。これは, 効用値を測定する2つの方法がそれぞれ独立した理論から考案されたことを反映している可能性がある。Brazier ら⁵²⁾によるレビューの中でも, 評点尺度法は時間得失法や基準的賭け法との相関は弱く, 評点尺度法の測定結果はむしろ, 痛みや身体機能といった健康状態との相関が高いことから評点尺度法は健康状態の一側面しか測定できないと述べている。さらに, 測定方法の信頼性については, 再テ

スト法による検討で時間得失法の信頼性が最も高かったと述べている。評点尺度法は効用値の測定方法としては理論的根拠が低く、EuroQol⁵³⁾でも補助的な手段として用いられており、妥当性、信頼性の上からは時間得失法による評価が適切と考えられた。

つぎに、評点尺度法と時間得失法によって効用値に差を認め、しかも時間得失法による効用値の値が評点尺度法のそれよりも低くなったことについて、以下に考察する。また、Hallanら²¹⁾によると、minor strokeの中間値が評点尺度法で0.71、時間得失法で0.88、major strokeはそれぞれ0.31、0.51となっている。いずれも、評点尺度法で求めた効用値よりも時間得失法で求めた効用値の方が高い値となっている。本邦では、効用値の算出に際して、評点尺度法と時間得失法の2方法の比較をしたものはこれまでのところ久繁らによる妊娠中毒症と先天性代謝障害についての報告⁷⁾とウィルソン病についての報告⁸⁾のみであるが、これによると、妊娠中毒症が評価尺度法で0.35、時間得失法で0.51、先天性代謝障害がそれぞれ0.53、0.62となった。ウィルソン病については、病型により両者の値がほぼ同様、もしくは評点尺度法で0.86、時間得失法で0.71などと時間得失法の値の方が低くなった。また、同様の研究としては、田村ら⁹⁾が看護婦を対象に難聴と車椅子状態について調査したものがある。これは基準的賭け法、時間得失法、人間得失法の3つの手法で効用値を測定したものであるが、難聴状態は先の方法から順に0.55、0.40、0.65、同様に車椅子状態はそれぞれ0.51、0.29、0.62となり、時間得失法で求めた効用値が他の方法で求めたものよりも低くなっている。Postらによれば、一般的に評点尺度法を用いて測定された効用値よりも時間得失法による効用値が高いとされている。Brazierら⁵²⁾も評点尺度法と時間得失法の結果からは両者に整合性のある関係は見出せないものの、大部分の報告で評点尺度法より時間得失法の方が高くなると報告している。今回は、対象者や測定方法などさまざまな要因が関連した結果、このような逆転現象が認められたと考えられ、今後検討していきべき課題としてとらえていきたい。

今回の研究の問題点と今後の課題について考える。まず、調査の対象者の中に脳卒中リスク患者

を含まなかった点である。Postらが述べたようにベッドサイドにおける意志決定に用いる効用値として最も適当とされる脳卒中リスク患者の効用値は本邦でも今後評価していく必要があると考えられる。

次に、調査方法についてである。本研究では一般的に理解が得にくい時間得失法に関して、補助的な手段も用いないまま、しかも、質問紙法で行った。本来は面接法で行われている³⁸⁾のものであり、対象者に質問の理解が得られたかどうかの疑問が残ってしまった。また、健康状態のシナリオについても不十分であり、今回の脳卒中の障害状態が対象者に正確な情報として理解されたとは言いがたい。健康状態のシナリオは一般的に二人称の散文で肯定・否定の両面を含み多面的に叙述されるべきと述べられている⁵⁴⁻⁵⁶⁾。このように調査方法については不十分な点が多く、今後改めていく必要があると考えている。

さらに、今回の質問では、死よりも悪い状態としての効用値の設定をしなかった。死よりも悪い効用値については、Patrickら⁵⁷⁾がマイナスの効用値で示すべきであると述べている。これについては、いくつかの健康状態についてカテゴリーに分類し、それぞれに効用値を設定したHealth Utilities Index⁵⁸⁾やEuroQol⁵³⁾などでもマイナスとして表示されている。臨床経済学で用いられる費用効用分析では、治療前の効用値と治療後の効用値の差を治療の結果とすることから、ある障害状態の効用値がマイナスになることはその差が大きく広がることになる。この問題については、調査の段階で十分な説明を入れることを前提に導入していく必要があると考えられた。

今後は、測定方法の信頼性を高め、基準的賭け法による検討も考慮した上で、我が国におけるさまざまな健康状態の選好を評価し、臨床経済学的分析に役立てていきたい。

(受付 2001. 5.31)
(採用 2002. 9.19)

文 献

- 1) La Puma J, Lawlor EF. Quality-adjusted life-years. JAMA 1990; 263: 2917-2921.
- 2) Robinson R. Cost-utility analysis. BMJ 1993; 307: 859-862.

- 3) Anell A, Norinder A. Health outcome measures used in cost-effectiveness studies: a review of original articles published between 1986 and 1996. *Health Policy* 2000; 51: 87-99.
- 4) Neumann P, Stone PW, Chapman RH, et al. The quality of reporting in published cost-utility analyses, 1976-1997. *Ann Intern Med* 2000; 132: 964-972.
- 5) Fanshel S, Bush JW. A health-status index and its application to health services outcomes. *Operations Research* 1970; 18: 1021-1066.
- 6) Gold MR, Siegel JE, Russell LB, et al. 医療の経済評価。(池上直己. 他監訳) 東京: 医学書院, 1999; 117-174.
- 7) 久繁哲徳. ハイリスク妊娠の危険度評価とその予測に関する効用分析の課題. ハイリスク妊娠に関する研究, 厚生省心身障害研究班, 平成3年度研究報告書, 1992; 94-102.
- 8) 久繁哲徳, 藤岡芳美, 佐々木暢彦, 他. スクリーニングの評価に関する研究. 新しいスクリーニングのあり方に関する研究, 厚生省心身障害研究班, 平成6年度研究報告書, 1995; 81-85.
- 9) 田村 誠, 野崎真実, 福田 敬. QALYsの効用値算出法に関する実証研究. *医療経済研究* 1996; 3: 87-103.
- 10) 二木 立. リハビリテーション医療のマクロ経済分析—「社会医療診療行為別調査」でみるリハビリテーション医療費. *総合リハ* 1994; 22: 593-601.
- 11) 二木 立. リハビリテーション医療の効率と費用効果分析—地域リハビリテーションを中心として. *総合リハ* 1994; 22: 1061-1068.
- 12) 二木 立. 日本の医療費. 東京: 医学書院, 1995; 173-197.
- 13) 渡邊 修, 宮野佐年, 船越政範, 他. 脳卒中リハビリテーション入院の費用便益分析—要介護度区分の視点から—. *臨床リハ* 2000; 9: 838-843.
- 14) 樋田美智子, 武村真治. 機能訓練教室の費用便益分析—仮想評定法によって測定された支払意思額を用いて—. *日本公衛誌* 2002; 49: 29-40.
- 15) 左手達男, 庄司豊彦, 三浦信義, 他. 大腿骨転子部骨折と医療費用効用分析. *整形外科* 1999; 50: 227-230.
- 16) 原田亜紀子, 川久保清, 李 廷秀, 他. 高血圧患者に対する運動療法の費用と効果に関する検討. *日本公衛誌* 2001; 48: 753-763.
- 17) 川久保清. 心臓リハビリテーションの費用・効果分析. *Heart View*; 1999; 3: 34-38.
- 18) Gore JM, Granger CB, Simoons ML, et al. Stroke after thrombolysis: mortality and functional outcomes in the GUSTO-1 trial: Global Use of Strategies to Open Occluded Coronary Arteries. *Circulation* 1995; 92: 2811-2818.
- 19) Gage BF, Cardinalli AB, Owens DK. The affect of stroke and stroke prophylaxis with aspirin or warfarin on quality of life. *Arch Intern Med* 1996; 156: 1829-1836.
- 20) Samsa GP, Matcher DB, Goldstein L, et al. Utilities for major stroke: Results from a survey of preferences among persons at increased risk for stroke. *Am Heart J* 1998; 136: 703-713.
- 21) Hallan S, Asberg A, Indredavik B, et al. Quality of life after cerebrovascular stroke: a systematic study of patients' preferences for different functional outcomes. *J Intern Med* 1999; 246: 309-316.
- 22) Duncan PW, Lai SM, Keighley J. Defining post-stroke recovery: implications for design and interpretation of drug trials. *Neuropharmacology* 2000; 39: 835-841.
- 23) Bosworth HB, Horner RD, Edwards LJ, et al. Depression and other determinants of values placed on current health state by stroke patients: evidence from the VA Acute Stroke (VAsT) study. *Stroke* 2000; 31: 2603-2609.
- 24) O'Meara JJ, McNutt RA, Evans AT, et al. A decision analysis of streptokinase plus heparin as compared with heparin alone for deep-vein thrombosis. *N Engl J Med* 1994; 330: 1864-1869.
- 25) Shin AY, Porter PJ, Wallace MC, et al. Quality of life of stroke in younger individuals: utility assessment in patients with arteriovenous malformations. *Stroke* 1997; 28: 2395-2399.
- 26) Lenert LA, Soetikno RM. Automated computer interviews to elicit utilities: potential applications in the treatment of deep venous thrombosis. *J Am Med Inform Assoc* 1997; 4: 49-56.
- 27) Adar R, Cohen E, Kreitler S. Carotid endarterectomy for symptom-free stenosis: the patient's point of view. *Cardiovasc Surg* 1994; 2: 582-585.
- 28) Solomon NA, Glick HA, Russo CJ, et al. Patient preferences for stroke outcomes. *Stroke* 1994; 25: 1721-1725.
- 29) Kwa VI, Limburg M, de Haan RJ. The role of cognitive impairment in the quality of life after ischaemic stroke. *J Neurol* 1996; 243: 599-604.
- 30) Mathias SD, Bates MM, Pasta DJ, et al. Use of the health utilities index with stroke patients and their caregivers. *Stroke* 1997; 28: 1888-1894.
- 31) Grootendorst P, Feeny D, Furlong W. Health Utilities Index Mark 3: evidence of construct validity for stroke and arthritis in a population health survey. *Med Care* 2000; 38: 290-299.
- 32) Dorman P, Dennis M, Sandercock P. Are the modified "simple questions" a valid and reliable measure of

- health related quality of life after stroke? *J Neurol Neurosurg Psychiatry* 2000; 69: 487-493.
- 33) Tengs TO, Yu M, Luistro E. Health-related quality of life after stroke: a comprehensive review. *Stroke* 2001; 32: 964-972.
- 34) Post PN, Stiggelbout AM, Wakker PP. The utility of health states after stroke: a systematic review of the literature. *Stroke* 2001; 32: 1425-1429.
- 35) Rankin J. Cerebral vascular accidents in patients over age 60: II. prognosis. *Scott Med J* 1957; 2: 200-215.
- 36) Patrick DL, Bush JW, Chen MM. Methods for measuring levels of well-being for a health status index. *Health Serv Res* 1973; 8:228-245.
- 37) Drummond MF, Stoddart GL, Torrance GW. 臨床経済学 (久繁哲徳, 西村周三監訳). 東京: 篠原出版, 1990; 127-168.
- 38) Bennett KJ, Torrance GW. Measuring health state preferences and utilities: rating scale, time trade-off and standard gamble techniques. *Quality of Life Pharmacoeconomics in Clinical Trials*, second edition B. Spilker, Lippincott-Raven Publishers, Philadelphia, 1996; 253-265.
- 39) Torrance GW, Thomas WH, Sackett DL. A utility maximization model for evaluation of health care programs. *Health Serv Res* 1972; 7: 118-133.
- 40) 久繁哲徳. 最新・医療経済学入門. 東京: 医学通信社, 1997; 123-124.
- 41) O'leary JF, Fairclough DL, Jankowski MK et al. Comparison of time-tradeoff utilities and rating scale values of cancer patients and their relatives: evidence for a possible plateau relationship. *Med Decis Making* 1995; 15: 132-137.
- 42) Salkeld G, Cameron ID, Cumming RG, et al. Quality of life related to fear of falling and hip fracture in older women: a time trade off study. *BMJ* 2000; 320: 341-346.
- 43) Mohide EA, Torrance GW, Streiner DL, et al. Measuring the wellbeing of family caregivers using the time trade-off technique. *J Clin Epidemiol* 1988; 41: 475-482.
- 44) 厚生省大臣官房統計情報部監修: 平成11年簡易生命表. 厚生統計協会, 2000.
- 45) 能登真一, 柳 久子, 戸村成男, 他. 脳卒中の障害状態についての効用値測定に関する再現性の検討—評価尺度法と時間得失法での比較—. 総合リハ 2001; 29: 357-360.
- 46) Froberg DG, Kane RL. Methodology for measuring health-state preferences—III: population and context effects. *J Clin Epidemiol* 1989; 42: 585-592.
- 47) Boyd NF, Sutherland HJ, Heasman KZ, et al. Whose utilities for decision analysis? *Med Decis Making* 1990; 10: 58-67.
- 48) Dolan P, Gudex C, Kind P, et al. The time trade-off method: results from a general population study. *Health Econ* 1996; 5: 141-154.
- 49) Froberg DG, Kane RL. Methodology for measuring health-state preferences—II: scaling methods. *J Clin Epidemiol* 1989; 42: 459-471.
- 50) Torrance GW. Social preferences for health states: an empirical evaluation of three measurement techniques. *Socio-Econ Plan Sci* 1976; 10: 129-136.
- 51) Churchill DN, Torrance GW, Taylor DW, et al. Measurement of quality of life in end-stage renal disease: the time trade-off approach. *Clin Invest Med* 1987; 10: 14-20.
- 52) Brazier J, Deverill M, Green C, et al. A review of the use of health status measures in economic evaluation. *Health Technol Assess* 1999; 3: 1-164.
- 53) EuroQOL group: EuroQOL. a new facility for the measurement of health-related quality of life. *Health Policy* 1990; 16: 199-208.
- 54) Llewellyn-Thomas H, Sutherland HJ, Tibshirani R, et al. Describing health states: methodologic issues in obtaining values for health states. *Med Care* 1984; 22: 543-552.
- 55) Gerard K, Dobson M, Hall J. Framing and labeling effects in health descriptions: quality adjusted life years for treatment of breast cancer. *J Clin Epidemiol* 1993; 46: 77-84.
- 56) Merlino LA, Bagchi I, Taylor TN, et al. Preferences for fractures and other glucocorticoid-associated adverse effects among rheumatoid arthritis patients. *Med Decis Making* 2001; 21: 122-132.
- 57) Patrick DL, Starks HE, Cain KC, et al. Measuring preferences for health states worse than death. *Med Decis Making* 1994; 14: 9-18.
- 58) Feeny DH, Torrance GW, Furlong WJ. Health utility index. *Quality of Life Pharmacoeconomics in Clinical Trials*, second edition, B. Spilker, Lippincott-Raven Publishers, Philadelphia, 1996; 239-252.

MEASURING UTILITIES FOR VARIOUS FUNCTIONAL OUTCOMES AFTER STROKE. COMPARISON OF RATING SCALE AND TIME TRADE-OFF METHODS

Shinichi NOTO*, Hisako YANAGI²*, and Shigeo TOMURA²*

Key words : stroke, utility, rating scale, time trade-off, QALYs, rehabilitation

Objectives To elicit utility of various functional outcomes after stroke for application in cost-utility analysis for stroke rehabilitation and to analyze the effects of demographic characteristics such as age, sex, professional status, marital status, or health status and the relationship between rating scale and time trade-off methods.

Methods The respondents were 460 healthy people, students (n=169), rehabilitation staff (n=134), family caregivers (n=52) and office workers (n=105). We assessed utility for 5 functional outcomes after stroke according to the Rankin scale, employing rating scale and time trade-off methods.

Results The mean utilities for Rankin I-V were 0.89-0.18 using rating scale, and 0.83-0.09 using time trade-off. Correlations between the two were only 0.176-0.412. The effect of professional status was clear, with rehabilitation staffs and caregivers giving higher utilities than students and office workers using time trade-off. However, effects of other characteristics were not apparent.

Conclusions People's preferences for stroke outcomes differ with disability level, and correlation between rating scale and time trade-off utility is weak. We found that few demographic characteristics had significant influence.

* Department of Occupational Therapy, School of Health Sciences, Niigata University of Health and Welfare

²* Institute of Community Medicine, University of Tsukuba