

## 連載

## 臨床経済学の基礎(12)

筑波大学大学院人間総合科学研究科 ヒューマン・ケア科学専攻  
保健医療政策学分野 教授 (社会医学系)

大久保一郎

第6回目の連載の時に効果の指標について解説を行なった。その時最も上位の効果の指標はQALY (Quality Adjusted Life Years: 質調整余命) であると述べた。そしてこのQALYは量の単位であるYOLS (Years of Life saved: 延長される生存年数) と、質の単位である効用値 (Utility) の積として、 $QALY = YOLS \times i$  ( $0 \leq i \leq 1$ ) で表される。より正確には毎年の効用値を生存年数で積分したものである (図1)。この図では④が最もQALYが大きく、最も余命の長い②ではない。今回はこの効用値 (i) の測定について解説する。

主な測定法には、Rating Scale Standard Gamble Time Trade-Offの3つがある (表)。測定したい状態について、本人の現在の状態でなければシナリオ等を作成して、被験者にその状態になったつもりになってもらう。そしてこれらの方法を用いて尋ね、回答を得るものである。

そもそも効用値は主観的なものであり、同じ状態であっても個人個人により異なる。例えば、喉頭がんの術後で声を出せない状態の効用値の測定を試みるとする。個人により効用値は異なるものであるから、職業が声優であれば、その効用値は他の人と比べて特に低くなることが予想される。逆に声を使わない職業であれば他の人より高いかも知れない。

この効用値は治療法を選択する上で、重要なこと

である。喉頭がんを例にとると、医師は患者に対して複数の異なる治療法を説明して、その選択肢の中から、患者の同意の下に一つの治療法を決定することとなる。これはまさしくインフォームド・コンセントである。過去においては医師が患者に十分な説明を行わずに、治療法を決定していたこともあったが、現在はそのようなことは許されない。もし、治療方法としては手術、放射線療法の2つが考えられるとする。医師は「手術の場合はこれから平均10年生きられるが、声は出ません。放射線の場合は声は出せますが、7年です。どちらを選択しますか。」と説明した。少し単純化しすぎているかも知れないが、要点は表していると思う。読者の皆さんはどちらを選択であろうか。手術を選ぶ人もいれば、放射線療法を選ぶ人もいて、どちらか一方のみとは考えられない。

この時、頭の中では  $QALY = YOLS \times i$  ( $0 \leq i \leq 1$ ) を計算しているのである。手術の  $QALY = 10 \times i$  (手術) と、放射線の  $QALY = 7 \times i$  (放射線療法) を比較して、大きい値となる方を選択するのである。単純に期待される生存年数が長い方を選択しているわけではないと思う。もし、患者が売れっ子のアナウンサーであれば放射線療法を、また作家であれば手術を選択する可能性が高いであろう。

このように効用値は同じ健康状態であっても、個人により異なるものである。だからこそ「説明と同意」が必要なのである。効用値を測定するには、事前に誰の効用値であるかを明確にする必要があ

図1 QALYの概念

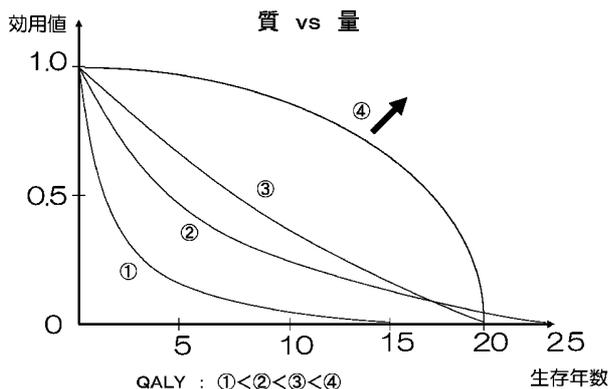


表 QALYの測定

- QALY = 延長される余命 × 1年あたりの質 (0-1)
- 1 : 100%健康な状態
- 0 : 死んでいる状態
- 1年あたりの質の主な測定方法
- ① Rating Scale
- ② Standard Gamble
- ③ Time Trade-off

り、研究目的に合った人の効用値を測定すべきである。その疾患を有した患者の効用値であるのか、それとも社会全体の効用値であるのか。後者であれば、国民の代表となるような被験者の選択をする必要がある。政策の決定や評価のために費用効果（効用）分析を行なうのであれば、個人や一定の限られた集団ではなく、社会全体での効用値の測定を行なう必要がある。

いずれにせよ、効用値の測定には複数の被験者の協力を必要とする。そして多くの場合、測定したい健康状態に対してシナリオを作成（複数の状態してはそれぞれ作成する）して、これに基づいて、各個人に尋ねることになる。以下3つの主な方法について解説する。

1. Rating Scale (図2)

まず、0から1、または0から100の目盛りの入った物差しを用意する。そして0は死んでいる状態で、1または100は完全に健康な状態であると説明する。次に測定したい状態を説明し、その状態は0から1、または0から100の物差しでどの程度であるかを指で示してもらおう。指された値が、その状態の効用値となる。例えば、介護度2の状態を尋ねて、0から100の物差しで50が示されたら、その人が考える介護度2の状態は0.5と測定される。

2. Standard Gamble (図3)

この方法は被験者にある賭け（ギャンブル）をする機会が与えられる。もしその賭けに勝ったら、そ

の状態（効用値  $i$ ）から脱して、完全な健康状態（効用値 1）に戻ることができるが、逆に負けてしまったら、その場で死んでしまう（効用値 0）ことになる。この賭けを選択するか否かは自身で判断できるというものである。賭けを行なうので、Standard Gamble（標準賭け法）という。

測定者は被験者にこの賭けに勝てる確率を伝え、賭けをするか否かを尋ねる。その確率が100%であれば誰もが賭けを選択し、逆に0%であれば誰も賭けをしない。つまり、この0から100の間に被験者が迷う確率、すなわち均衡がとれる確率が存在するはずである。この値から測定したい効用値（ $i$ ）を求めることができる。図2から分かるようにこの確率が50%であれば、 $i = 1 \times 0.5 + 0 \times 0.5 = 0.5$ となるのである。

3. Time Trade-Off (図4)

この方法では、測定したい状態が一定期間継続すると仮定し、その状態と100%健康な状態のより短い期間とを交換するものである。例えば、ある健康状態（効用値  $x$ ）が10年継続し、その後死亡するものとする。この状態と100%健康な状態と交換することができる。当然100%健康な状態の期間が10年であれば、誰でも交換するであろう。一方、この期間が0年であれば誰でも交換しないであろう。つまり、この0から10年の間に被験者が迷う年数、つま

図2 Rating Scale

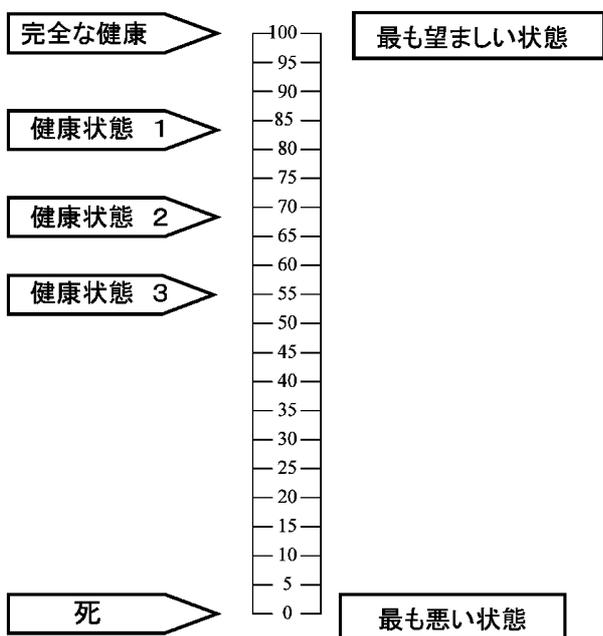


図3 Standard Gamble

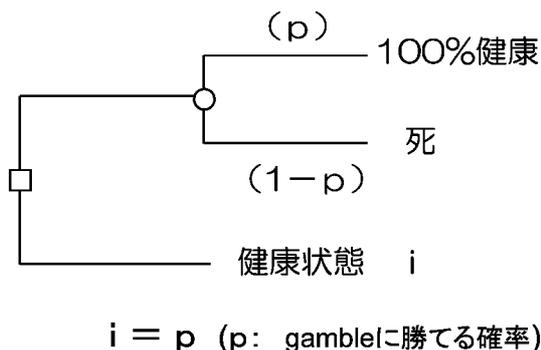
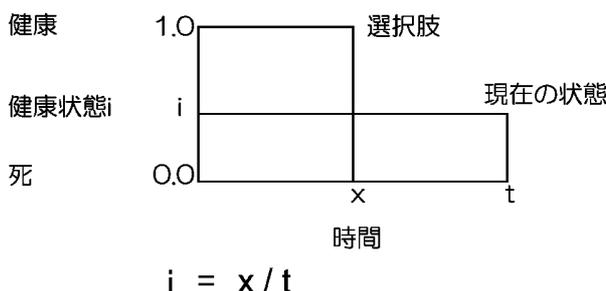


図4 Time Trade-off



り均衡する年数があることになる。あまり良くない健康状態で10年過ごすのであれば、多少寿命を縮めても100%健康な状態を獲得したいと思うのである。例えば、5年で均衡が取れるとすれば、 $i \times 10 = 1 \times 5$ つまり  $i = 0.5$ となる。時間との交換をするので、Time Trade-Off (時間得失法) という。

#### 4. 3つの測定方法の短所と長所

Rating Scale は測定方法としては、比較的簡単に被験者からも理解を得られやすい。また、同時に複数の状態を一度にその物差しで測定することができる長所もある。ただし、他の2つの方法と比較して、正しく測定できないとされていて、一般により低い値が測定される傾向にある。また一方、効用値そのものより、複数の状態の順位付けを行なうのは、正しく簡便に測定できるという特徴も有している。効用値そのものより、順位付けに向いているため、他の2つの測定方法と合わせて実施されることが多い。また、1人の人の日々変化する健康状態が、悪化しているのか改善しているかの傾向を把握

するには適していると思われる。

Standard Gamble 法は理論的に3つの方法の中では、最も正しく効用値を測定できるとされている。それは健康状態との交換を一定の確率で可能とするが、その確率が医療の世界における不確実性を反映しているからである。しかし、3つの方法の中では、確率という概念そのものが理解しづらいこともあり、最も方法的には難しいとされている。

Time Trade-Off 法は3つの中では最も頻繁に利用されていると思われる。質問用紙にも落とすことができ、被験者にも理解されやすい。しかし、その交換が確実に行われるという点で、理論的には標準賭け方式より劣るとされている。

倫理的な視点からは、Standard Gamble 法も Time Trade-Off 法も「死」との取引となる。前者は賭けに負ければその場で「死」となり、死か健康かという選択を迫ることになる。後者は寿命を短縮させてまでも、より健康な状態を獲得する取引を行うことになる。そのため、倫理的に不快を感じる人も多いであろう。