健康関連 QOL : もうひとつのアウトカム指標

京都大学大学院医学研究科健康解析学 福原 俊一

Health-Related Quality of Life : An Overview

Shunichi FUKUHARA

Kyoto University, Department of Epidemiology and Health Care Research,
Graduate School of Medicine and Public Health
Yoshidakonoe-cho, Sakyo-ku, Kyoto 606-8501, Japan

〈Abstract〉

Under the prevailing "disease-oriented" paradigm of health, Quality of Life (QOL) has been treated with profound skepticism. Two facts, however, militate for taking QOL seriously into account: first, both social and clinical decision making today require evidence presented in the form of patient outcomes; second, of patient outcomes, one of the most important from the patient's standpoint is QOL. Giving several examples, the author discusses why traditional endpoints are insufficient and why patient-based outcomes are relevant in modern social and clinical decision making.

And yet not all outcomes are relevant: for the purposes of assessment of health and medical care, QOL should be confined to its health-related elements, and instruments to measure health-related quality of life (HRQOL) should be rigorously tested with psychometric methods. Even more important, however, is the instrument's content validity: i.e., whether it measures what one wants to measure. This paper lists criteria for selection of appropriate instruments and describes the recent debate and moves to formulate guidelines on the use of HRQOL evidence in pharmaceutical product approvals and manufacturers' product efficacy claims. Finally, the author discusses the future of HRQOL research and uses of HRQOL in clinical practice.

Key words : Quality of Life, outcomes research, patient-based outcomes, clinical trial

はじめに

クオリティオブライフ (以下 QOL) を key word とした医学論文が増加している。医学論文全体の数が増えているので相対的判断が必要となるが、1967 年に最初の論文が出版されて以来、絶対数としては確実に増えており、図 1 に示すように、これは世界、アジア、日本に共通しており、当該分野への関心の高さを物語っている。では、その質はどうだろうか？

Gill や中山らは、QOL を key word にした内外の医学研究の質について批判的な文献レビューを行った。Gill らは、QOL を key word とした医学研究論文の中で、概念的に QOL を明確に定義していたものは 15%、測定する対象を明確に規定したのは 47% にすぎないことを報告している¹⁾。中山らは、国内文献レビューを行っており、「QOL の領域の規定、評価尺度の選択などが十分に選定されていない場合が多く、分析的な研究対象として

図 1 Quality of Lifeをキーワードとした医学論文 (Medline)
(1966-1999, All Language)

表 1 QOL への懐疑

- ・客観性に乏しい
- ・QOL の定義が統一されていない
- ・QOL データが医師や患者を mislead する
- ・QOL は、医療介入以外の多くの要因によって影響を受ける
- ・有効性、安全性以上の情報は得られない(臨床試験)
- ・症状や副作用と同じことであるので不要
- ・QOL データの結果の臨床的意義や解釈が不明確である

表 2 患者立脚型アウトカム

- ・健康関連 QOL
- ・選好に基づいた尺度 (preference-based index)
- ・患者満足度
- ・医療介護負担度

よりもスローガンやキャッチフレーズとして扱われやすい傾向がある」ことを指摘している²⁾。このような点から考えると、質はまだまだのようである。

このように多くの QOL 研究に質が伴っていないこともあってか、医学界で QOL 研究が手放しで受け入れられているわけではない。患者の QOL を尊重することは、確かに総論、スローガンとしては受け入れられるが、各論になるとためらいや抵抗がみられる。事実、医学研究で QOL を用いることへの懐疑は依然として大きい。とくに生命医学を中心に発展してきたわが国の医学研究のパラダイムのもとでは、医学研究に QOL を用いることへの抵抗はさらに大きいと考えられる。表 1 に、QOL に対してよく聞かれる懐疑的な意見を示す。

では、QOL に関するこのような懐疑にもかかわらず、QOL を用いた評価研究をこれほどまでに後押ししている原動力は何だろうか。この背景に関しては、次節(1節)でより詳しく述べたい。とくにこの 20 年間に推進されてきたアウトカム研究³⁾における位置付け、さらに EBM の時代の日常診療に QOL データがどのように活用されるかについても触れる。

たとえ背景や必要性は理解できたとしても、QOL 研究に対する懐疑にすべての回答が得られるわけではない。これらすべてに反論を試みることは本稿の意図ではないが、実際の医学研究で QOL を取り上げる前に明らかにしなければならない事項は多い。たとえば、表 1 に挙げた中でも懐疑の中核を占める部分が、「QOL の定義が一定していない」「QOL など測れるのか?」ということがあるが、2 節では、QOL とは何か、測定方法には科学性があるのか、などについて述べながら、これらの懐疑について答えたい。

³⁾ アウトカム研究とは：医療評価には、構造、プロセス、転帰（アウトカム）の 3 つの要素を評価することが必要とされているが、おもにアウトカムに着目して行う研究の総称。縦断的観察研究や介入研究のデザインをとることが多い。

本誌は薬剤評価に関する雑誌であるので、この領域における QOL に対する最近の動向についても 3 節で触れたい。最後に QOL 研究に関する今後の課題を概説する。

1. QOL を用いた評価研究を後押ししている原動力は何か？

QOL 研究に対する懐疑にもかかわらず、QOL を用いた医学研究が盛んに行われるようになった背景は、健康・医療領域において、「現代の社会レベルでの決定 (social decision making) に QOL データが必要とされているから」「現代の診療上の決定 (clinical decision making) に QOL データが必要とされているから」の 2 つに集約されるのではないかと考える。前者はアウトカム研究や医療経済評価を中心としたヘルスサービス研究の流れにつながり、後者は EBM に基づく日常診療の流れにつながっているといえよう。社会の高齢化、慢性疾患がほとんどを占めるようになった疾患分布の変化、これに伴う医療の目標の変化と多様化、国民・患者の自律性指向の高まりなどは QOL 重視の背景としてよく言われることではあるが、つきつめれば、上記の 2 点に集約されるのではないだろうか。

1) 社会レベルでの決定 (social decision making) に QOL データが必要とされている

Social decision making に有用な情報を与える研究の 1 つとして、アウトカム研究、ヘルスサービス研究は、北米を中心にこの 20 年間急速な発展を遂げてきた³⁾。これらの研究は、人口の高齢化、慢性疾患の急増、テクノロジーの急速な進歩などに伴う医療需要の急増や医療の高額化の中にあつて、有限の医療資源をいかに有効に配分するかという今日の社会的状況が求めていることであつた。

また近年のアウトカム研究における最大の特徴は、従来の疫学研究で用いられなかった評価指標が積極的に取り上げられるようになったことである。なかでも際立っているのが、患者立脚型アウトカム (patient-based outcomes) である。表 2 にその代表的な指標を示す。

このような研究では、医療社会学、行動科学、計量心理学、医療経済学など、これまでの疫学研究では必要とされなかった多分野の研究者によって研究チームを構成する必要が生じた。北米にはもともとこれらの研究分野に質の高い人材の層が厚かったことも、これを容易にした。

アウトカム研究という新しい研究の流れが形成されるとともに、QOL は患者立脚型アウトカムの 1 つとして明確に認識されるようになった。ここでは、QOL はもはや曖昧な概念しか持たないスローガンではなく、測定目的と測定する構成概念が明確に規定され、計量心理学的手法により検証済みの尺度を用いて定量的に測定することが可能で、統計的に処理されうるアウトカム変数の 1 つとしてみなされるようになった。スタンダードな尺度も開発され、このうちのいくつかはさまざまなアウトカム研究で広範に使われるようになっていく。たとえば、MOS 36-Item Short-Form Health Survey (SF-36) は現在、世界の多くの臨床試験や疫学研究等で用いられている⁴⁻⁶⁾。

また peer review を要する一流の医学雑誌にも、患者立脚型アウトカムを用いた論文が多く見られるようになった。たとえば、Lee Goldman らが行ったペースメーカー治療の効果を QOL で評価した研究は、雑誌 “New England Journal of Medicine” の巻頭論文として発表されている。この研究では、従来型のペーシング治療方法が、生理学的には優れるとされるより高価なペーシング治療法と比較して、QOL の点での改善効果に差を認めないという結果を示し、患者の QOL を考慮しつつ社会的に効率的な治療法の選択に貢献した⁷⁾。

経済評価をはじめとする研究も盛んに行われてきた。オーストラリア、カナダ、そして欧州諸国の一部では、新薬をはじめとする新規の医療テクノロジー適用の承認、価格の決定において、その効果と効率性を検討した結果を提出することが規制当局によって求められている。米国においては、規制当局ではないが、マネージドケアなどの保険支払い側がこのようなデータを求めている。このような経済評価のためには、患者の価値観に照ら

し、特定の健康状態の質を評価したQOL単一指標（選好に基づくQOL尺度で測定）が、必須となっている（詳細は池田論文p.99～105を参照）。また前途のSF-36からもこの単一指標を得る試みを、筆者は英国のBrazierと共同して研究を開始した。

これまで行われた一連のアウトカム研究から、QOLデータの社会レベルでのもう1つの有用性を示唆する結果が示されている。これは、ベースラインで測定したQOLデータが、将来の病休、医療資源消費、さらに生命予後までも予測することである。QOLを構成する基本的な要素（2節で述べる）の中では、「身体機能」が予測性を有するという報告が最も多く、またこの結果は直感的にも容易に理解できる⁸⁾。しかし、このことは、「身体機能」に限定されていない。たとえば「全体的健康観（General Health Perception）」が、いくつかのQOL構成要素の中でも最も強い予測性を示したという報告が複数ある⁹⁾。さらに「メンタルヘルス」が、透析患者のような有病者においても一般住民においても、同様な関連性や予測性を持つという結果が示されている¹⁰⁾。このような研究成果は、患者立脚型アウトカムを用いた地域レベルあるいは産業レベルのhealth monitoring（その典型が健診事業）やdisease management（疾患を包括的に管理するための新しい概念）、さらにこれらのプログラムの評価等に活用されることを期待させる。

2) 患者個人レベルでの治療法の選択・決断のために患者立脚型アウトカムが必要とされている：日常診療におけるQOL

患者の生活の視点に立脚したアウトカムを用いた研究からうまれたevidenceが、治療法の選択や決断に重要である理由を以下に述べる。

〈現代の治療選択において、単独のアウトカムでは不十分である〉

現在、医療の大半を占める慢性疾患の治療においては、治療法がただ1つであることは少なく、互いに競合する複数の選択肢があることがほとんどである。たとえば早期前立腺癌の治療では、根治手術、放射線療法、内分泌療法、さらに「何も

しない様子を見る（watchful waiting）」という選択肢まである。このような場合、個々の選択肢の有効性と安全性を、生存などの客観的なエンドポイントを用いて、ランダム化比較試験（RCT）などでそれぞれの治療オプションの有効性と安全性を科学的に検証したり互いを比較することは可能で、事実、信頼できるevidenceが得られていることも少なくない。しかし、複数の治療法から最適な治療を選択・決定する患者にとって、これらのevidenceだけでは十分ではないのである。たとえば前立腺癌は高齢者に多い疾患だが、彼らにとって生命予後の延長はすべてではない。むしろ、「尿漏れ」「排便障害」「性機能の低下」などによるQOLの低下こそが、より重大な関心事であることが多いくらいである。実際の治療選択において、このようなデータがrelevantな情報として求められているのである。この例に限らず、現代の多くの慢性疾患における治療選択・決断のために、患者の生活の視点に立脚したアウトカムを用いたevidenceが今後活用されることが予測される。

〈医師と患者の評価の乖離〉

臨床医は患者の最大利益のために最も適切であると考えられる検査や治療を行っているわけであるが、だからといって医師が患者の状態をいつも適切に評価しているとは限らない。とくに疾患が患者の主観的健康度や日常生活機能にどのような影響を与えているかについて、そうであることが報告されている。たとえばUCLAのLitwinは、前立腺癌手術後の患者の症状や主観的な健康度・日常生活機能の評価について、患者の主治医と患者自身の間でどのような違いがあるか報告した¹¹⁾。この報告では、患者が認識している健康度や日常生活の影響を、医師が過小評価している可能性を指摘している。医師と患者の視点の間にあるこのような乖離を、QOLを用いたevidenceがすべて埋めてくれるとは考えられないが、日常診療において少なくとも補助的な役割をもつことが今後期待される。4節の今後の展望の中でも触れたい。

〈症状とQOLの違い〉

症状は患者の主観的な健康評価の最たるもので

あり、あえてQOLのような曖昧な指標をさらに用いることはない、という意見もよく聞かれる。しかし、症状とQOLのデータには大きな違いがある。

- 症状が患者の主観的な特性であることは間違いないが、報告されるためには、少なくとも2つのステップが必要となる。まず患者が医師にこのことを伝えなければいけない。伝えたとしても、これをみた医師が症状として有意であると認めカルテに記載する必要がある（患者は自分の健康状態の異変に気づいた場合でも医師にこれを訴えない割合がかなり高いこと、さらにたとえ患者が訴えたとしても、医師がそれを有意と認めないことが少なくないこと、結果的に患者が感じた健康上の異変の1~2割ぐらいしかカルテに記載されないこと、などの報告がある）。
- 症状はそれぞれ単独に扱われる。たとえば前立腺癌手術直後の患者によく見られる症状として、排尿困難と尿もれがある。症状の頻度や程度に関しては、これを定量化する「症状スケール」で排尿困難、尿もれの程度をそれぞれ定量的に測定することは可能である。しかし、排尿困難と尿もれが同時に存在する場合、そのどちらが患者の日常生活により大きなインパクトを有しているかは、症状スケールだけで明らかにすることはできない。症状ではないが、疾患の活動性や重症度を反映するとされてきた生理学あるいは生化学的検査の結果が、患者の主観的な健康度と関連しないという報告¹²⁾さえもいくつかある。
- 「メンタルヘルス」は、健康関連QOL(定義については2節参照)を構成する最も基本的で重要な要素であり、ほとんどの健康関連QOL尺度に含まれている。一般外来を訪れる患者で、身体的、生理的、生化学的に何の異常も認めない割合は30%とも80%とも言われている¹³⁾。

メンタルヘルスは、脳内の生理学的・生化学的なメカニズムも関連しており、すべてがfunctionalあるいはbehavioralな問題とは言えないのは事実であるが、現在得られる診断テクノロ

ジーでは、その診断は医師による病歴と診察が中心を占めている。メンタルヘルスの専門家以外の医師が、多忙な外来の中で患者のメンタルヘルスを適切に評価することは至難である。この点で、メンタルヘルスを定量的に測定する健康関連QOL尺度が活用されることが期待される。これは外来に限らず、産業保健や地域保健にも敷衍できることである。

以上述べたように、日常診療レベルにおいても、患者の生活視点にたったQOLのデータを用いた研究のevidenceは、患者の価値観や視点を尊重した診療選択を行うための有用な情報を提供することから、その有用性が期待されている。

2. QOL とは何を測るのか？

QOL は測れるのか？

冒頭表1にQOLへの懐疑を示したが、これに一部答えるかたちをとりながらQOLの概念や測定について述べたい。

1) 「QOLの定義が曖昧」

QOLは主観指標であるので、「QOL」という実態があるわけではなく、またあまりにも多くの意味を持ちうるため、確かにスローガンのようになってしまう傾向が強い。事実、多種多様な概念でQOLが定義されている。しかし、どのような概念・要素としてのQOLを測定するかは、研究目的によって規定されることを再認識する必要がある。逆にいえば、研究目的を明確にすることによってのみ、測定概念を特定することができるのである。「QOLを測定した、QOLが改善した」というような漠然とした表現がよくみうけられるが、正確には「QOLを構成すると研究者が考えた」特定の概念を測定しているのである。研究目的を明確にし、そのためにQOLを構成する「何を」測定したいのかを特定することによって、「曖昧である」などの批判や誤解を避けることができると考える。

「定義の曖昧さ」の批判のもう1つの重要な背景として、QOL研究や、QOL尺度の開発過程の多くにおいて、測定しようとするQOL、およびこれと他の変数（要因、アウトカム）との関連を示す概念モデルが構築される必要があるにも関わらず、

図2 健康関連QOL
(Health-Related QOL)¹⁵⁾

これが構築されていないことがある。QOLを用いた医学研究において、この概念モデルの構築には、2つの異質なパラダイムを持った研究者、すなわち疾患モデルに依拠している臨床研究者と社会学を専門とする研究者が協力する（あるいは社会学の素養を持った臨床研究者が単独で）必要がある。Wilsonらは、この2つのパラダイムに基づく概念モデルを統合する新たな健康関連QOLの概念モデルを提唱している¹⁴⁾。わが国の医学研究領域に医療社会学の研究者が少ないことも、QOL研究の進展を遅らせている一因であると考えられる。

2) Overall QOL vs 健康関連QOL

薬剤評価のように、QOLをアウトカムあるいはエンドポイントの1つとして医療評価のために用いる際には、測定要素はさらに明確に限定される。この目的で使用する時、医療介入以外の外的な要因に大きく影響されるいわゆる overall QOL（生きがい、経済状態、ソーシャルサポート、満足度）ではなく、健康に起因し医療介入によって改善できる可能性のある領域に測定範囲を限定する必要がある。これを、「健康関連QOL」「Health-related QOL」（以下 HQOL）と呼ぶことが、国際的にもコンセンサスがほぼ得られつつある（図2）¹⁵⁾。

QOLを構成する最も基本的な構成要素に関しても、国際的なコンセンサスはできつつある。1つは「身体機能」で、階段を昇れるか、独力で服を着られるというような内容。2番目は「心の健康」，「メンタルヘルス」で、とくに気分の落ち込みや不安が含まれる。3番目は「社会生活機能」で、これは一般的な社会関係というよりも、身体機能やメンタルヘルスの変化によってもたらされた友人関

表3 適切なQOL尺度を選択する基準

- ・表面的妥当性を有しているか？ さらに、その尺度に含まれる項目は回答者に理解しやすいか？ 理解できても回答しにくいということはないか？ (feasibility, acceptability)
- ・自分の臨床研究の結論を当てはめたい集団 (target population) に対して尺度が検証されているか？
- ・翻訳された尺度の場合、言語的に、内容的に、概念的に原版と一致しているか？
- ・尺度の構成概念妥当性、内容的妥当性は十分か？ (概念の構成は、研究目的、測定目的に合致しているか？ 尺度の項目の内容は、自分の測定したい内容を測っているか？)
- ・信頼性、下位尺度の内的一貫性は適切か？
- ・その結果を臨床的、社会的な面からの解釈に生かすことができるか？ そのための基礎データがあるか？
- ・その結果を診療における治療法の選択や医療政策的な判断に生かすことができるか？
- ・既存の尺度に比べてどのような優れた点があるのか？

係やつき合いの変化を指すことが多い。これらの3つの要素は、いみじくもWHOの健康に関する定義ともほぼ同じである（最近、WHOの定義の4番目の要素として霊的な面を加えることが議論されているようであるが、本稿では触れない）。加えて、健康状態に起因する「日常役割機能」（仕事や家事）も最も基本的なQOLの要素として含まれることが多い。さらに、痛み、活力、睡眠、食事や性生活なども重要な要素として含まれることがある。

現在QOLを用いた医学研究で見られる大きな問題の1つは、overall QOLを測定するための項目がHQOLを測定するための項目と混在している尺度を用いて医療介入を評価している例が少なからず見られることであり、大きな混乱を招いている。このような尺度を用いて介入評価を行っても、当然反応性の低い結果しか得られないというようなことが起こる。目的に最適な尺度を選択することの重要性がここにある。表3に尺度の選択基準を示す。

さらに、同じHQOLといっても、測定のレベルがいくつかに分かれることを認識しておく必要がある。「階段を2階まで上ることがむずかしい

か?」「落ち込んでいるか?」といった function-based あるいは status-based なレベル,「現在の健康状態はかなり悪い」「現在の健康状態は私の日常生活にかなりの負担を与えている」というような evaluative なレベル,そして「現在の健康状態に満足している」「現在の状態は死より悪い」というような価値付けをしたレベルの3段階に大別できよう。HQOL を測定する場合,これらのどのレベルで測定しようとしているかを明確に認識する必要がある,この異なるレベルでの測定を一緒にすることによる混乱も避ける必要がある。したがって,機能,日常生活への負担 (burden),満足度,効用値はそれぞれ独立に測定する必要がある。

3)「QOLは測れるのか?」

まず「QOL データは客観性に乏しい」という批判があるが,これは的外れとも言えるものである。なぜなら,QOL は主観指標であるからである。よく「客観的な QOL」などの表現が見られるが,これは矛盾した表現である。しかしこの「客観性」には,「測定方法に信頼性があるのか?」「科学性があるのか?」といった意味合いも含まれていると考えられ,これに対しては一定の手続きを経た測定方法や尺度を用いることを示す必要がある。HQOL は主観指標であるため,これを測定する尺度を開発するときは,主観特性の測定に関する「お作法」に従う必要がある。本稿では詳述を省くが(詳細は大橋論文 p.71~82 を参照),これはおもに計量心理学(古典的テスト理論)によっているものが多い。現在臨床系の医学雑誌においても,このような一定の手続きを踏まない方法で QOL を測定した研究は,すでに受け入れられないようになっている。このような厳密な手続きにより一定の基準をクリアした尺度を用いることによって,HQOL という主観特性が測定可能であること,そしてこれを活用した HQOL 研究に対する評価も高まることが期待される。

しかし,HQOL を測定する尺度の信頼性,妥当性は,尺度の性能を示す重要な評価項目ではあるものの,これだけでは不十分である。最も肝要なことは,前述した「何のために」という目的と,

「何を測るか」という「中身」の問題である。これは,「内容的妥当性」の問題といえる。QOL のどのような概念・要素を測定したいのかということを経験者は明確に意識し,質問項目のそれぞれを厳密にチェックすることから尺度開発が始められなければならない。

4) QOL の特徴

以上 QOL に対する懐疑へ答えるかたちで HQOL の定義や測定について述べた。さまざまな批判がある一方で,HQOL データにはいくつかの画期的な特徴がある。HQOL の最大の特徴は,HQOL の測定が患者の視点に立脚して行われることにある。ADL も HQOL に近いように見えるが,決定的な違いは ADL が第三者の観察者を介して測定されるのに対して,HQOL は患者の健康度やこれに起因する日常生活機能を患者の目を通して報告させているところにある。QOL 尺度は QOL の多面性を考え,多次元尺度で測定されることが多い。しかしここでいう多次元とは,医療提供側が使用している臓器別に分類によったものではない。あくまでも日常生活機能(たとえば役割機能,社会生活機能など)によって分類されたもので,それぞれの機能を構成する「下位尺度」に含まれる項目は,医学的な用語ではなく,徹底して「患者の日常生活活動の言葉」(例:階段の昇降,仕事・家事,普段の付き合いなど)によって表現されることになる。その他にも,1)患者が自分の健康度を直接報告する。医師の解釈を経ない。2)病気を持つ人も健康人も区別せず,健康度を連続的な概念として測定する。患者だけでなく,症状を呈していない患者や健康人の健康度とそれに関連した生活の質を測定することによって,患者との比較も可能となる。3)健康がどの程度阻害されているかというネガティブな側面(dysfunction)だけでなく,どの程度健康か(well-being)というポジティブな側面をも測定する,などの特徴がある。

3. 薬剤評価における患者立脚アウトカムの活用 —最近の動向

臨床試験のための一連の ICH 指針に含まれる

表 4 QOL データが有用である例・根拠

- ・生命予後や臨床的な有効性ではとらえられない効果に関する付加的な情報を提供する
- ・臨床的な有効性や生命予後では差が認められないときに有用
- ・多くの慢性疾患に対する医療評価に有用。とくに、形態学的、生化学的には効果を捉えられない疾患で有用。
- ・個々の症状（陽性、陰性効果）の程度を症状スケールなどでいくら定量化したとしても、どの症状が患者の日常生活、社会生活にどのようなインパクトを与えているかが評価できない。
- ・医師と患者の間で、症状や生活機能の評価に乖離があることが少なくない。
- ・ベースラインのQOLが、将来の生命予後、医療資源の消費（外来利用度、入院日数、医療費など）の重要な予測因子であることが種々の研究で明らかにされている。

「一般指針」は、筆者も expert working group の一員として参加して作成された。これには、「臨床試験は、よくデザインされたプロトコールのもとに行われなければならない」と明記されており、このプロトコールのもとに行われて得られた結果を、適切な統計学的手法を用いて有効性と安全性の面から科学的に評価が行われることになる。現在、すべての臨床試験でこれが徹底されているわけではないが、以前に比較すれば格段の進歩が見られていることは歓迎すべきことである。

しかし、無作為に割り付けられた対照群と比較して、客観的な指標で測定したアウトカムが統計的に有意差がある（ない）ことを示すだけで充分だろうか？ 近年、このことを疑問視する指摘が見られるようになってきている。事実、QOLをはじめとした患者立脚型アウトカム指標を用いる新薬の臨床試験が増加している。その背景としては、前節で述べたことがほとんど同様にあてはまる。すなわち、臨床試験において、生命予後や臨床指標の改善が依然として重要なエンドポイントではあるが、それだけでは十分ではないとの認識が高まってきているのである。患者立脚型アウトカムを用いたデータが、ある種の新薬の承認や価格決定において、また実際の診療場面における治療薬の選択にあたって、このようなデータの必要性が医師からも患者からも求められるようになってき

表 5 臨床試験における QOL 評価に必要な要件

1. スタディデザイン：臨床試験全般に要求される要件と同じ
2. HQOL をエンドポイントとする正当性、仮説
3. 特定の HQOL 尺度を選択する根拠、尺度の特性
 - ・ HQOL のどの要素・領域（domain）を測定するのか？
 - ・ その QOL 尺度や測定する domain は研究の目的・仮説と合致しているか？
 - ・ その QOL 尺度は標的集団の特性に合致しているか？（ceiling や floor 効果の程度）
 - ・ 尺度の信頼性、妥当性、反応性の検討結果の明記
4. HQOL 測定方法のプロトコール明記
 - ・ 測定手段（質問紙、面接、電子的）
 - ・ 測定時期
 - ・ 欠測値を最小限にする対策
5. 結果の解釈：統計的、臨床的、社会的
6. HQOL 解析計画のプロトコールへの明記
 - ・ スコア方法
 - ・ 欠測値の扱い
 - ・ 検定方法、多重性への対策
 - ・ Effect size の設定（MCID）
 - ・ パワー計算、サンプル数計算

たのである。また、ICH 以前にわが国独自の判定基準として、担当医が薬剤の有効性と安全性を総合的に評価するところの有用性という指標があった。HQOL は、治療による陽性・陰性効果を、総合的にしかも患者の視点で定量化するという特徴があるため、一種の「患者の視点からみた有用性」と考えることもできるかもしれない。

米国 FDA は、これまで QOL を臨床試験のエンドポイントとして用いることに消極的な姿勢をとってきた。その理由としては、「有効性と安全性以外のデータは必要ない」を始め、表 1 に示した QOL に対する懐疑と重複するところが多い。Gotay らは抗悪性腫瘍薬の臨床試験で QOL がエンドポイントとして適切である場合があるとして、いくつかの例をあげている¹⁰⁾。表 4 に、QOL がエンドポイントとして有効な例・根拠を、Gotay 文献に著者の意見も交えたものを示す。筆者も、QOL をすべての臨床試験に用いる必要はなく、ケースバイケースで判断すべきであるという

FDA の Temple らの見解に同意する点が少なくない。たとえば、腎性貧血のエンドポイントとして従来血色素が用いられてきたが、近年 HQOL のほうが患者、社会にとってより relevant な指標であると考えられるようになっている¹⁷⁾。ところが、急性心筋梗塞や肺炎のような急性疾患の短期的な効果に関しては、QOL は relevant な指標とはいえず、依然として重症化や生命予後こそが最も重要な指標であろう。しかし、たとえ生命予後に有意な差が明らかであっても、なお QOL の使用が重要性をもつ場合がある。患者の中には生命予後より QOL のほうを重視するものもいるはずで、治療選択の際に重要な情報となるからである。

現在、薬剤評価と QOL をめぐる主な論点は、QOL への効果を添付文書 (labeling) と宣伝 (claim) に使用することが許されるか、許されるとしたらどのような基準が適切かということである。現在、欧米の企業と規制当局の間で、この点に関するガイドライン作りの作業が行われている。表 5 に現在提案されているガイドラインの骨子を示す。

わが国においては、臨床試験における QOL データの位置付け、取り扱いなどについての議論がまだ活発に行われていない。もちろんこれに関するガイドラインづくりは行われていないが、早晚、必要になると予測される。産、官、学による検討が期待される。

4. 今後の展望

1) EBM の実践と QOL

QOL 研究の成果は、EBM (Evidence-based medicine) の 4 段階といわれる実践のステップのうち、その最終段階にあたる「EBM を実際の患者に適用する」という最も重要な場面で、今後ますます活用されることが予想される。すなわち、EBM に基づいた診療では最終的な診断・治療方針の決断を医師と患者が協力して行う (shared decision making) 必要がある。そしてこれを行っていくためには、患者の視点にとって重要性のある情報 (evidence) を提供する必要がある。すでに

北米では、このような治療選択・決断を補助するビデオや CD-ROM などの媒体が開発され、実際の診療にも普及しつつあり (decision support system), これらの媒体の多くでは、QOL のような患者立脚型アウトカムを用いた evidence が提供されている¹⁸⁾。この shared decision making にあたっては、患者—医師間のよいコミュニケーションが必要であり、そのためには QOL の情報を共有することが役立つ。そもそも慢性疾患の治療では、医師が主役となって病気を治すというよりは、患者自身の疾患への理解、受容、治療へのコンプライアンスが必須であり、この意味で治療は医師と患者との共同作業といえる。また、患者の生活の視点に立脚した evidence を、医療提供側にフィードバックすることにより、医師は患者の状態を単に医学的なモデルだけでなく、bio-psycho-social なモデルの中で理解することができる。その結果、これまで医師が気付かなかったような患者への影響に気が付き、より良いコミュニケーションを促し、患者—医師関係や患者の治療へのコンプライアンスを改善することによって患者の QOL が良好に保たれる、というフィードバックループが作られることが期待されている¹⁸⁾。わが国には、まだこのような考え方は定着しておらず、したがって必要なインフラも整備されていない。また、わが国の患者自身がはたしてこのような意思決定方法を望んでいるかも不明である¹⁹⁾。この領域に関する研究の発展が切に望まれる。

2) 結果の解釈

統計的有意差が出たとしても、その差にどのような意味があるかの結果の解釈が重要であるという考え方は、何も QOL に限ったことではない。解釈にあたっては、distribution-based interpretation, anchor-based interpretation, population-based interpretation などの方法がとられる。たとえば anchor-based interpretation は、あるエンドポイントの差が他の外的な基準、たとえば臨床的あるいは社会的な基準などに照らして、どのような意味があるかの解釈を行うことである。臨床的に意味のある最小差, minimally clinically

表 6 薬剤評価における HQOL に関する指針 (案)

- ・スタディデザインの明示 (目的, 内容, 尺度を用いた測定の実際, ブラインド化, 欠損値の最小化の方法など)
- ・尺度開発 (患者からの情報のもとに項目は開発されたか, 項目選択の方法, 尺度のどの領域が使われるのか, スコア方法とその根拠)
- ・尺度の validation (信頼性, 内的一貫性, 妥当性, 反応性, 等の検証)
- ・解析計画, 結果の解釈の明示 (事前に設定した effect size のみに基づく方法, 臨床指標など他の評価尺度による結果との関連による臨床的意義づけ)
- ・薬剤の添付文書や商品宣伝における表現内容ではできるだけ specific に: (すべての結果の公開, 「QOL」のようなあいまいな表現を避け測定した特定の領域 (domain) を明示)

meaningful difference (MCID) の概念は客観的なエンドポイントについても使われ, 試験開始前のサンプルサイズ推定のための重要な決定因子の1つである²⁰⁾. これは, 臨床試験のプロトコルに求められる要件の1つでもある (表6). 結果の解釈という問題は, QOL では客観的なエンドポイントよりさらに重要である. たとえば QOL 尺度で測定した結果 5 点差を認めた場合, 統計的に有意であってもこれにどのような意味があるかは不明であることがほとんどだからである. したがって, これらの解釈に必要なデータを提供する種々の研究が必要となる. これらの研究こそが, まさにアウトカム研究やヘルスサービス研究の典型ということになる. このようなデータがなければ, QOL を用いた臨床試験などで MCID を事前に設定することは困難である. わが国には, データの得られにくさ等諸種の理由でこの領域の研究が遅れており, 今後発展が期待される.

3) より精度の高い尺度へ

2 節で, QOL を用いた研究の evidence が日常診療での治療選択に役立つことを述べた. これをさらに 1 歩進めて, 個々の患者さんから得たデータを患者本人に還元し, 臨床に応用できないものか? という質問と要望は, 少なからず聞かれる. しかし, 現存の少数項目 (回答者への負担を考慮し, 50 以下の項目数のものが多い) の QOL 尺度では, データ精度の問題のため, アウトカム研究の

表 7 QOL 研究の課題

- ・QOL データの解釈に必要なデータを提供する研究: その QOL のスコア差は, 臨床的に, 社会的にどのような意味を持つのか?
- ・文化的なレスポンスバイアスへの対処
- ・レスポンスシフトバイアスへの配慮
- ・個人レベルでの応用が可能な測定方法の開発
- ・価値付けデータを用いた social decision making の倫理性

ような群レベルでの研究には役立つが, 個人へのフィードバックはかなり困難であると考えられている²¹⁾. しかし, このような問題点を克服するために, 近年新しい手法を活用した測定方法の開発・研究が進んでいる. これは項目応答理論を活用して, コンピュータ上で adaptive な測定を行うもので, すでに北米では資格試験などに応用され, 実用化されている. この方法を使うことにより, プールされた多数の項目のうち, 最小限の数の至的な質問項目を自動的に選択し患者に答えてもらうことにより, 高い精度の測定を, しかも患者に負担を強いずに行うことを可能にするのである. これらの研究の展開次第によっては, 将来的に QOL 測定の臨床現場における個人応用が可能になることも期待される²²⁾.

その他, 表7に示すごとく, QOL 研究に残されている課題は山積みしている. 今後の研究の発展に期待したい.

文 献

- 1) Gill T, Feinstein A. A critical appraisal of the quality of quality-of-life measurements. *JAMA* 1994; **272**(8): 619-26.
- 2) 中山健夫, 田中平三, 野地有子. わが国における「QOL 研究」の動向—文献のレビュー—. 第2回日本行動医学会抄録集 1995; 20-1.
- 3) Tarlov AR, Ware JE, Greenfield S, et al. The medical outcomes study, an application of methods for monitoring the results of medical care. *JAMA* 1989; **262**: 925-30.
- 4) Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36), I: conceptual framework and item selection. *Med Care* 1992; **30**: 473-83.
- 5) Fukuhara S, Bito S, Green J, Hsiao A, Kurokawa

- K. Translation, adaptation, and validation of the SF-36 Health Survey for use in Japan. *J Clin Epidemiol* 1998 ; **51**(11) : 1037-44.
- 6) 福原俊一 : MOS Short Form 36 items Health Survey—新しい健康アウトカム指標, 厚生 の指標 1999 ; **46**(4) : 40-5.
- 7) Lamas. GA, Orav EJ, Stambler BS, et al. Quality of life and clinical outcomes in elderly patients treated with ventricular pacing as compared with dual-chamber pacing. *N Engl J Med* 1998 ; **338**(16) : 1097-104
- 8) Tsuji I, Kuwahara A, Nishino Y, et al. Medical cost for disability : a prospective observation of National Health Insurance beneficiaries in Japan. *J Am Geriatr Soc* 1999 ; **47** : 470-6.
- 9) Wilson I, Kdler EL, Kasl S. Health perceptions and survival : do global evaluations of health status really predict mortality? *J Gerontol* 1991 ; **46** : S 55-S 65.
- 10) Young EW, Meredith DT, McCullough A, et al. Hospital readmissions and length of stay : The Dialysis Outcomes and Practice Pattern Study (DOPPS). *J Am Soc Nephrol* 2000 ; **11** : 248 A, A 1306.
- 11) Litwin MS, Lubeck DP, Henning JM, Carroll PR. Differences in urologist and patient assessments of health related quality of life in men with prostate cancer. Results of the CaPSURE database. *J Urol* 1998 ; **159** : 1988-92.
- 12) Alonso J, Anto JM, et al. Measurement of general health status of non-oxygen-dependent chronic obstructive pulmonary disease patients. *Med Care* 1992 ; **30** : MS 125-MS 135.
- 13) Barsky AJ, Cleary PD, Sarnie MK, et al. The course of transient hypochondriasis. *Am J Psychol* 1993 ; **150** : 484-8.
- 14) Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. *JAMA* 1995 ; **273**(1) : 59-65.
- 15) 福原俊一. いまなぜ QOL か—患者立脚型アウトカムとしての位置づけ, 臨床家のための QOL 評価ハンドブック. 医学書院, 2001 (印刷中).
- 16) Gotay, CG, et al. Quality-of-life assessment in cancer treatment protocols : reserach issues in protocol development. *J Natl Cancer Inst* 1992 ; **84** : 575-79.
- 17) Canadian Erythropoietin Study Group. Association between recombinant human erythropoietin and quality of life and exercise capacity of patients receiving haemodialysis. *BMJ* 1990 ; **300** : 573-8.
- 18) Inui TS, 福原俊一. EBM マインドを育てる. *EBM ジャーナル* 2001 年 1 月 : 90-102.
- 19) Matsumura S, Bito S, Liu HH, Fukuhara S, Yamamoto K, Wenger NS. Acculturation of end of life decision making attitudes : Cross-cultural survey in Nagoya and Los Angeles. *J Gen Intern Med (supple 2)* 1999 ; **14** : 53
- 20) Jaeschke R, et al. Measurement of health status—ascertaining the minimally clinically important difference. *Control Clin Trials* 1998 ; **10** : 407-15.
- 21) McHorney CA, et al. Individual-patient monitoring in clinical practice : Are available health-status survey adequate? *Qual Life Res* ; **4** : 293-307.
- 22) Ware JE, Bjorner JB, Kosinski M. Practical implication of item response theory and computerized adaptive testing ; a brief summary of ongoing studies of widely used headache impact scales. *Med Care* 2000 ; **38**(9 suppl) : 1173-82.