

Quality of Life 研究の在り方とその現状

田崎 美弥子 (東京理科大学理学部教養)

Quality of Life 研究は、がん患者のホスピス活動から患者本位の医療の在り方を模索したことから始まったものである。しかし、1985年に北米の Food and Drug Administration (FDA) が新薬承認の際に同じ薬効があれば QOL の向上を一つの指標としたことで、治験の評価試験関連の論文が急に増加した。2001年の2月の時点で、Medline データベースで Quality of Life を研究課題とする論文数は、1983年の562件から、その約10倍の5622件に達している。

QOL 研究はすでに20年近くたっているものの、90年代初頭に WHO が指摘した医療場面における患者の健康指標に関する問題点は依然として残されているように思われる。当時、患者の健康評価指標として、生存率や有病率、日常生活動作や行動へのインパクト測定を目的とした Sickness Impact Profile (Bergner, Bobbitt, Carter et al, 1981)、患者の健康認識指標として開発された Nottingham Health Profile (Hunt, McKenna and et al, 1993)、障害や機能状態計測が目的の the MOS SF-36 (Ware et al, 1993) が存在したものの、どれも疾患による障害や影響に焦点をあてたものであり、そのため、WHO は、疾患や症状の消滅だけに焦点を当てるのではなく、患者の良好状態というポジティブな面を含んだ健康指標を開発するために WHOQOL 調査票開発プロジェクトを開始した。また、その頃すでに30近い QOL 調査票が存在していたが、殆ど北米、イギリスで開発されているため、文化的な偏りは否めず、真の意味での国際的な調査票を開発するために、開発途上国を含んだ世界各国の研究者で同時期に、討議を繰り返すことで、文化・言語的等価性を重視した研究プロトコルを用いた。この調査票開発に9年の歳月をかけ、現在100項目の研究版と26項目からなる臨床用に開発された短縮版が発表されている。

日本においても1991年に厚生省の「抗悪性腫瘍薬の臨床評価に関するガイドライン」において、第III相試験に抗腫瘍効果と生存期間に加えてエンドポイントとして QOL 評価を加えたことで、QOL 研究は90半ばにはブームといえるような活況を呈した。実際、筆者らが、QOL 論文のレビューを目的として、1995年に日本科学技術情報センターの医中誌国内医学文献ファイルおよび医学中央雑誌刊行会のデータベースに、Quality of Life, QOL, 生活の質、クオリティ・オブ・ライフ、ヒト、人間、症例報告を入れ検索したところ、キーワードとして用いているものは、5000件以上が検出したが、QOL 調査を実施していたものは、200件であり、カテゴリー別に見ていくと、1) 悪性腫瘍および他の慢性疾患患者の QOL 評価、2) 薬物や術式の評価指標としての QOL 利用、3) 心理的課題、生物学的な指標、および時間の推移などと QOL との検討に分類された。が、そのうち、1) QOL が明確に定義されている、2) 調査票の原典(オリジナルか翻訳版

か)が明記されている、3) 調査票の開発過程が明記されている、4) 統計的手法が採用されている、という基準に合致したものは20にも満たなかった。現在、徐々に優れたQOL評価調査が報告されるようになってきたものの、依然として健康指標の在り方についての統一見解がないため、QOL調査には大きな見解の相違があると思われる。

ここでは、一定水準に達するために、北米のMedical Outcome Trustが設定しているInstrument Review Criteriaを適用し、QOL調査の現状とその問題点を挙げ、本来のQOL調査の在り方を考察したい。Instrument Review Criteriaは、1) 概念モデルや測定モデル、2) 信頼性 3) 妥当性、4) 反応度、5) 解釈度、6) 負担、7) 代替的様式、8) 文化・言語からのみた適用性の8項目から構成されている。

1) 概念モデルや測定モデル

日本におけるQOL調査に示されているQOL観には大きく分けて、主観的指標と客観的指標の2つの視点があるという立場を取るものと、主観的指標のみとする立場を取るものがある。しかし、多くのQOL調査においては、ADLと混同した論旨のものがあり、少し精神面を含んで、バランスがとれているように見えるものも、開発過程で身体症状に関する項目の比重が増し、考察にいたるとその身体症状にほぼ限定するような形でQOL評価を行っているものが多いのが特徴である。一方、WHOQOLではQuality of Lifeを「個人が生活する文化や価値観のなかで、目標や期待、基準、または関心に関連した自分自身の人生の状況に関する認識」と定義し、身体的、心理的、自立のレベル、社会との関係、信念、環境の重要な側面との関わりとして広範囲な概念であり、文化、社会、環境における良い面、悪い面の両側面を含んだ主観的な評価に重点をおいている。

2) 信頼性

内的整合性として一般的にCronbach's Alphaの最小値が0.7であること、他の調査票との相関係数が0.75であること、再テスト再現性をみるための分散分析などの結果が付記されていることが奨められているが、記述統計がある場合は、付記されているが、逆に一切の記述が無い場合に分けられる。日本で筆者らが行った一連のWHOQOL調査では、Cronbach's Alphaが0.87~0.97にあり、スコアが高ければ、身体精神度が悪くなるGeneral Health Questionnaire(GHQ)との相関係数が、-0.45~-0.47であり、身体精神度が悪ければQOLが低くなる関係が見られた。

2) 妥当性

妥当性には、内容の適切さ、尺度の明瞭さ、理解しやすさといった内容的妥当性、その概念にかかわる理論的意味合いにおいて、調査票のスコアがいかに正しく得られるかといった概念関連妥当性や、調査票のスコアが基準的な測定法といかなる関係にあるかを示す基準関連妥当性や、その結果からみた将来の予測的妥当性が挙げられる。内容、概念妥当性を示すためには、調査票の開発過程の情報と、データ収集方法の記述が明らかにされている必要がある。基準関連妥当性に関しては、健康状態の評価に関する広く受けられたこと基準測定法がないので検討されることはない。

WHOQOL調査票では、約2年間の歳月をかけて、開発プロトコルについての討議を進め、質的調査、予備調査、本調査、フィールド調査の過程を踏んでいる。その過程については毎回公表されて、質問項目があらゆる文化に適合するように、フィードバックがかけられた。日本における QOL 調査では、殆どが、欧米ですでに開発された調査票をそのまま翻訳あるいは少し改訂していたが、そのような調査票では、回答率の低い項目があり、それには原文の翻訳上の問題も絡んでいると推定されるものもある。生活様式、文化の異なる地域で開発された調査票を日本で使用する際の問題点が浮き彫りにされた。

4) 反応度

同一標本について介入を行う場合、その前後のスコアの変化について記述されている必要があること、調査票を用いた場合は、統計解析に必要な標本数であること、臨床比較試験では無作為抽出をするのが前提であり、対照群の有無、異なる治療法を用いる方法、術式の効果を同定するために対象患者の同質性が統計的に検定されている必要がある。残念ながら、標本数が揃っていても、無作為抽出法は臨床試験では特殊な疾病患者を収集することが困難なため、外来に来る患者全てを対象にすることは、QOL 調査のみではなく、治験では比較的一般的なようである。対象患者の同質性についても、触れている論文のほうに限られている。WHOQOL 調査の中では、標準化を計るために、東京、大阪、長崎の 1410 名を電話帳からの層別化無作為抽出法を用いて郵送調査を実施したが、各地域での年齢、性別による有意な差異が認められなかった。

5) 解釈度

量的なスコアを臨床場面や一般的に理解されている意味を持つ質的な分類に言い換えることである。特に QOL 値の差異がどの程度臨床的な意味があるかについては今後の課題になると思われる。

6) 負担度

負担度には、回答者が、調査票を記入する際にかかる労力、時間、心理的負担すべてを含む回答者負担と、調査者がその調査票を準備、処理するための平均時間の両者が含まれるが、調査によっては、3種類から4種類の調査票を実施しているものもあり、調査側の意図が前面に出て、回答者負担についてあまり考慮されていないものも見受けられる。あまりに多くの調査票を用いた場合、当然回答者の飽きから欠損値が多くなるなどのバイアスがかかると思われるので、その点も考慮すべきであろう。また、欠損値の中には、宗教や性的活動などの質問項目など、心理的抵抗があるために、あえて回答を拒否する場合もあり、質問項目の構成にも配慮が必要だと思われる。

7) 代替的な様式

自計自記式、他計自記式、熟練観察者による評価、コンピュータ補助自記式、達成測定、代理記入版といった様々な様式がある時に、代替的な様式を用いた場合、比較可能であるかどうかのポイントになる。例えば、実際の場面では、高齢者を対象にした場合、調査票の文字の大きさなどがポイントになり、調査者が読み上げて代理記入する場合も

あるが、そういった場合、スタイルの違いについて特に差異がないかどうかの検定も必要であれば行うべきだと思われる。

8) 文化・言語の適用性

WHOQOL 調査票は、異文化間の言語における等価性を得るために、各地域で、言葉の言い回しや、概念定義のわかりやすさを専門家、非専門家からなるグループで質的調査を行った上で、翻訳・逆翻訳を繰り返したが、その結果、わかりづらい日本語になるなど弊害が生じたが、国際比較を念頭においた場合、それは避け難いことであるかもしれない。しかし、そのような煩雑な手続きを行って開発された調査票は上記のように非常に限定されている。

残念ながら日本においては QOL 研究の偏在化は顕著であるように思われる。これは概念の吟味が十分にしつくされないまま、臨床的な必要性（特に新薬開発上の要請）から急速に現場に導入されてしまったからであろう。もともと客観性や科学性を最重視してきた分野であり、患者自身による主観的評価という微妙なデータ研究の枠組み取込むことの難しさがあったためと思われる。実際、QOL 評価のあいまいさが、医療場面での生理学的評価と同等に扱うことができないと感じる医療関係者は多いと思われる。しかし、WHO の健康定義が「単なる疾病や衰弱の欠如ではなく、心理的、身体的、社会的良好状態」であるのならば、QOL 評価を医療の現場で行うことの意義は、医療の受け手が感じる疾病による影響とともに医療サービスの質を、医療の担い手が把握し、支援するためのものであり、医療サービスの目的は、疾病だけを治すのではなく、その疾病を持つ人、そのものの心身の健康に伴う幸福度、物質的な幸福度、社会人間関係における幸福度などを向上することであると理解し、そのために必要な受け手の主観的評価に対する認識が必要であろう。