投稿

医師仕事関連QOL測定ツールの開発

江面（エヅラ） 美祐紀＊1＊2　澤田（サワダ） 克彦＊3　宅島（タクシマ） 祐介（ユウスケ）＊4
滕（テン） 麗達（リダ）＊5　五十嵐（イガラシ） 中（アタル）＊6＊7

目的　2020年以降のコロナ禍で，医療従事者側の負担軽減も価値の要素として認識されるようになり，医療従事者の負担を可視化する観点から，「主観的」なアウトカム尺度の整備が望まれる。また，昨今の医療技術の開発においてもこれまでの概念が覆され享受できる価値を多面的に評価する取り組みが必要とされる。本研究では，医療従事者のうち患者や医療技術と直接に接する機会の多い医師に焦点をあて，仕事関連の価値を構成する指標としてQOLを直接測定できる尺度を開発することを目的とした。

方法　QOL尺度の標準的な開発方法を参考に，医師を対象に４ステップで調査をした。まず，医師20名に対し影響因子を収集するために定性的なインタビューを実施し1stバージョン質問票を作成した。次に８名の医師にバリデーションインタビューを実施し2ndバージョン質問票として修正した。さらに374名の定量調査で妥当性および信頼性の検証を実施した。その結果をもとに3rdバージョン質問票として修正し，確認調査として，その信頼性と基準関連妥当性のための調査を実施した。また，QOL測定の専門家を交えた検討会を実施し，各種の妥当性を検討・実証しながら質問票を構築した。

結果　医師の仕事関連のQOLの影響因子から，仕事量，仕事時間，協力，診療，労働条件，労働環境，心理面，ワークライフバランス，キャリアの９項目が導き出され，それぞれの項目ごとに質問を作成し検証した結果，Cronbach α係数は0.899であった。

結論　開発された医師の仕事関連のQOLを測定するための質問票は，オリジナルのプロファイル型尺度で，９項目５段階で構成され，単純合計でスコア化して使用できるものとなった。

キーワード　仕事，医師，QOL，測定，質問票

Ⅰ 緒　　言

コロナ禍で医療資源のひっ迫が議論の的となり，患者や家族のみならず医療従事者側の負担軽減も価値の要素として認識されるようになった。2020年に実施されたアンケート[1)]でも，医療負荷・医療従事者の負担の軽減を重要な要素として挙げた回答者の割合は，家族の負担や労働生産性の割合よりも高かった。また，昨今の技術開発において薬剤領域ではバイオ医薬品や再生医療技術等の導入など，医療機器においてはロボット手術やアプリをはじめとするDigital

＊1 大塚ホールディングス株式会社経営企画部医療経済担当顧問
＊2 慶應義塾大学大学院健康マネジメント研究科特任准教授
＊3 大塚メディカルデバイス株式会社臨床開発部ディレクター（統計解析）
＊4 大鵬薬品工業株式会社メディカルアフェアーズ部課長補佐
＊5 東京大学大学院薬学系研究科医薬政策学特任助教　＊6 同客員准教授
＊7 横浜市立大学医学群データサイエンス研究科准教授

Transformationの導入など，開発の視野が以前に増して広くなり，その技術力で改善される範囲も主観的・客観的に広くなっている。医療従事者側の負担を可視化する際に，施術時間や入院期間のような「資源消費量（resource use）」の観点以外から医療従事者自身の負担を評価した研究は限られている[2]。QOLが患者報告アウトカム（PRO）[3]の流れから研究が進められたことも一因であるが，医療従事者の負担を可視化する「主観的」なアウトカム尺度の整備が望まれる。先行して実施したシステマティックレビュー（以下，SLR）[2]では，今回のテーマを十分にカバーできる既存の尺度が存在しないことが明らかになった。以上より，PROの医療従事者版，本研究では仕事関連の価値を構成する指標として医療従事者のうち患者や医療技術と直接に接する機会の多い医師のQOLを直接測定できるプロファイル型尺度を構築することを目的とした。

Ⅱ 研 究 方 法

QOL尺度の標準的な開発方法[4]を参考に４ステップで質問票の開発を進め，調査会社データベースに登録されている医師から全国統計分布[5]で調整した医師を対象とした。なお，本研究における調査は質問票開発を目的とした内容の特定，検証のための調査であり個人の行動や意識に関する情報収集を目的としたものではないため倫理委員会の審査は受けていない。調査にあたっては，目的・調査協力の任意性・調査結果を他の目的には利用しないこと，個人情報は特定できないことなどの説明を行ったうえで同意を取得し実施した。

（１） 定性インタビュー（第１ステップ）

医師の仕事関連のQOLに影響する因子を広く収集することを目的に，医師20名に対し[6]質的調査として，イントロダクション（経験年数，管理職，スタッフ数など），自由回答から関連因子を収集する純粋想起および仮説因子（患者，人間関係，病院との関係，新技術，診療報酬，労働環境・条件）を提示しながら収集する助成想起から構成されたインタビューガイドに従い，１名につき約１時間のインタビューを実施した。インタビュー内容のテキスト分析（KH Corder）に基づき質問項目を絞り込み，1stバージョン（以下，1st）質問票を作成した。

（２） バリデーションインタビュー（第２ステップ）

1st質問票を使用し，表現の適切性，解釈可能性，質問票の改善点等の検証を目的に，医師８名に対し１名につき約１時間のインタビューを実施し，その結果を受け2ndバージョン（以下，2nd）質問票に修正した。

（３） 定量調査（第３ステップ）

2nd質問票を使い妥当性と信頼性の評価を目的に374名にWeb調査を実施し，JMP（Ver.16, SAS Institute Inc.）を用いてデータを解析した。結果等について２名の専門家と検討会を実施し，3rdバージョン（以下，3rd）質問票に修正した。

（４） 確認調査（第４ステップ）

3rd質問票および基準関連妥当性を評価するために，visual analogue scale（VAS）を用いて「総合的な仕事関連QOL」の追加質問を用い，98例を対象に[7]Web調査を実施，JMPで解析をし，専門家検討会実施後に質問票を最終化した。

Ⅲ 研 究 結 果

2020年３月から2022年７月にかけて４ステップで実施した調査および質問票の作成結果を段階ごとに示す。

（１） 定性インタビュー結果

1） テキスト分析

インタビュー中出現頻度の高い単語はQOL，時間，患者，仕事，ストレス，手術であった。純粋想起と助成想起各々においてポジティブ（＋）とネガティブ（－）な回答語句を特定し，いずれかの方向（＋または－）ならびに双方向

性を示した語句を回答内容の強度別（強調／普通）に**表１**にまとめた。純粋想起から抽出された要素を総合して「ワークライフバランス（以下，WLB)」の項目を構築し，助成想起からは，「治療」「時間」「労働環境」「インフラストラクチャー（以下，インフラ)」「キャリア」「協力」の項目を構築した。

2） 項目抽出

1）で得られた項目に，先に実施したSLR[2]から「心理」を加え，「時間」「治療」「インフラ・労働環境」「協力」「WLB」「キャリア」「心理」の７項目を定めた。

3） 1st質問票

７項目について項目網羅性と内容妥当性を検討し，複数の質問文を作成した。回答形式は段階評定法を用い，項目の特性により「ゼロ（あるいはほとんどない)」から「非常（あるいは最大)」までの段階を測定する単極性の形容詞尺度，ある考え方について是認する方向とその逆の２方向に測定可能な双極性のリッカート尺度を用いた５段階とした[4]。

（２） バリデーションインタビュー結果

1） 表現の適切性

項目ごとに検討した結果，「時間」は勤務時間と仕事量に分けること，「協力」では“本来の仕事”の意味が曖昧であること，“同僚”は範囲が狭く“まわり”という表現が適していること，「治療」では“うまくいっている”という表現が治療結果を指すと誤解されることが明

表1　マインドマップ分析結果のまとめ

カテゴリー	単方向（＋）因子	強調[1]	普通[2]	単方向（－）因子	強調[1]	普通[2]	双方向（＋／－）因子		強調[1]	普通[2]	構築項目
純粋想起	治療・手術成功 自由時間等 給料・手当 自己裁量	0 3 1 1	1 6 2 0	当直・オンコール 研修医教育 雑務・雑用 幹部の方針	2 0 1 1	4 2 1 2					WLB
助成想起											
患者	治療・手術成功	4	9	突発的対応・急患 無駄話 患者家族対応 患者理解力	1 1 2 1	6 2 5 12	患者数多さ 難しい治療	＋ － ＋ －	1 1 0 0	3 5 2 3	治療 時間
労働環境労働条件	医師・スタッフ数 自由時間等 施設の充実度 振休 通勤時間 手術機会	5 1 1 1 1 0	2 3 5 0 4 3	当直・オンコール カンファレンス 研修医教育	3 0 0	3 5 1	研究・学会	＋ －	4 0	8 6	労働環境 インフラ
病院との関係	施設評判 給料・手当 自己裁量	1 2 0	3 9 2	幹部の方針 他施設競合 他科業務 雑務・雑用	1 1 1 2	1 0 0 2	昇進	＋ －	0 0	3 1	キャリア インフラ
人間関係							医師 看護師 MR スタッフ	＋ － ＋ － ＋ － ＋ －	1 1 0 0 0 0 7 1	3 12 1 5 3 2 10 2	協力
診療報酬							診療報酬	＋ －	1 1	6 5	協力／時間
新技術							新技術	＋ －	3 0	10 7	治療／時間

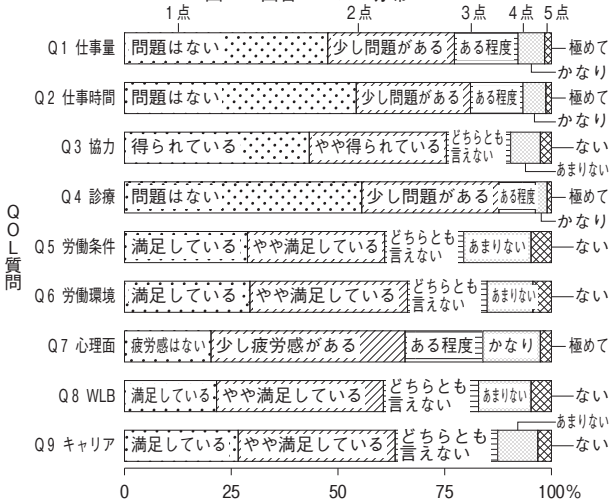
注　1）　強調は，“とても”“非常に”“特に”などの修飾語が使われた回答
　　2）　普通は，修飾語を用いない回答

らかになった。「心理」では“疲れ”が精神的・肉体的どちらを指すか不明確であり，「インフラ・労働環境」では労働条件と労働環境を分けたほうがよいことが示唆された。

2） 解釈可能性

項目ごとの解釈について，「時間」では能力を聞かれていると意図しない解釈をした例があった。“本来の仕事”“診療や治療”では解釈が一律ではなかった。「心理」での“疲れ”は，精神的・肉体的双方の疲労感や精神的のみと解釈された。

図１　回答スコアの分布



3） 回答形式

回答の選択形式の違いによる検討では，「治療」で単極性，双極性両方で同一回答であったが，その他の項目において単極性よりも双極性の方が低いスコアになる傾向があった。回答の選好は時間と協力が単極性であった以外すべて双極性の回答形式であった。

4） 質問票全体について

質問数，回答選択肢（５段階）については妥当であるとの意見が大半であった。

5） 2nd質問票

質問票は９つの質問で回答は１～５までの選択肢で再構成した。Q1は仕事量，Q2は本来の仕事に費やせる時間（以下，仕事時間）として分割させ，Q3，Q4，Q7は文言を変更し，Q5を「労働条件」，Q6に「労働環境」として分割させた。また，内容的妥当性を確認するために各項目に新たな検証質問（Q10～Q28）を作成し対応させた。

（３） 定量調査結果

1） 回答結果分析

Q1からQ9の回答は選択肢４番目，５番目の分布割合が低かったが，Q4は特にその傾向が強かった（図１）。

表２　各回答と検証質問の相関係数[1)]

質問項目	検証質問[2)]（上段：カテゴリ　下段：質問番号）											
	患者数[3)]	本来業務[4)]	協力状況[5)]	診療状況	患者関係	勤務時間[6)]	給与	設備体制	機器・薬剤	強い疲れ[7)]	仕事外の時間	キャリア
	Q10 Q11	Q13 Q14	Q15 Q16	Q17	Q18	Q19 Q20	Q21	Q22	Q23	Q24 Q25	Q27	Q28
Q1 仕事量	**0.322**	0.213	0.135	0.346	0.181	0.407	0.355	0.275	0.109	0.540	0.478	0.278
Q2 仕事時間	0.300	**0.157**	0.206	0.278	0.231	0.391	0.346	0.312	0.167	0.425	0.433	0.171
Q3 協力	0.157	0.094	**0.306**	0.380	0.137	0.143	0.231	0.318	0.214	0.316	0.285	0.276
Q4 診療	0.208	0.166	0.073	**0.422**	**0.276**	0.204	0.261	0.350	0.227	0.415	0.269	0.324
Q5 労働条件	0.345	0.189	0.101	0.413	0.180	**0.358**	**0.569**	0.304	0.189	0.477	0.417	0.396
Q6 労働環境	0.307	0.130	0.192	0.384	0.182	0.273	0.473	**0.477**	**0.317**	0.467	0.410	0.395
Q7 心理面	0.321	0.179	0.110	0.394	0.205	0.332	0.385	0.318	0.180	**0.661**	0.426	0.299
Q8 WLB	0.329	0.220	0.147	0.408	0.221	0.386	0.433	0.383	0.224	0.508	**0.603**	0.388
Q9 キャリア	0.207	0.075	0.105	0.404	0.153	0.272	0.452	0.355	0.271	0.327	0.299	**0.661**

注　1）Spearman相関係数を絶対値で示した。
2）複数の質問で構成された検証質問は各々の質問回答に基づいた統合スコアを導出し，Spearman相関係数算出に用いた。
3）Q11回答値÷Q10回答値
4）Q14回答値÷Q13回答値
5）Q15，Q16の回答内容に基づく合成変数を導出
6）Q20の回答値÷Q19の回答値
7）Q24の回答値＋Q25の回答値
8）各質問項目に対応する検証質問の相関係数を太字で示した。

図2　医療従事者（医師）の仕事関連のQOL調査票（a）およびスコア計算に用いた値（b）

（a）医療従事者（医師）の仕事関連のQOL調査票 直近1カ月のことを想像して回答してください。		（b）スコア計算に用いた値（点）
Q1	あなたの「仕事量」についてもっともあてはまるものを選んでください。	
	□　仕事量には問題はない	5
	□　仕事量には少し問題がある	4
	□　仕事量にはある程度問題がある	3
	□　仕事量にはかなり問題がある	2
	□　仕事量には極めて問題がある	1
Q2	あなたが「本来の仕事に費やせる時間」についてもっともあてはまるものを選んでください。	
	□　本来の仕事に費やせる時間には問題はない	5
	□　本来の仕事に費やせる時間には少し問題がある	4
	□　本来の仕事に費やせる時間にはある程度問題がある	3
	□　本来の仕事に費やせる時間にはかなり問題がある	2
	□　本来の仕事に費やせる時間には極めて問題がある	1
Q3	あなたの「同僚（周り）からの協力」についてもっともあてはまるものを選んでください。	
	□　同僚（周り）から協力を得られている	5
	□　同僚（周り）からやや協力を得られている	4
	□　同僚（周り）から協力を得られているかどうか，どちらとも言えない	3
	□　同僚（周り）からあまり協力を得られていない	2
	□　同僚（周り）から協力を得られていない	1
Q4	あなたの「診療」についてもっともあてはまるものを選んでください。	
	□　診療に問題はない	5
	□　診療にあまり問題はない	4
	□　診療に問題があるかどうか，どちらとも言えない	3
	□　診療に少し問題がある	2
	□　診療に問題がある	1
Q5	あなたの「労働条件」についてもっともあてはまるものを選んでください。	
	□　労働条件に満足している	5
	□　労働条件にやや満足している	4
	□　労働条件に満足しているかどうか，どちらとも言えない	3
	□　労働条件にあまり満足していない	2
	□　労働条件に満足していない	1
Q6	あなたの「労働環境」についてもっともあてはまるものを選んでください。	
	□　労働環境に満足している	5
	□　労働環境にやや満足している	4
	□　労働環境に満足しているかどうか，どちらとも言えない	3
	□　労働環境にあまり満足していない	2
	□　労働環境に満足していない	1
Q7	あなたの「疲労感」についてもっともあてはまるものを選んでください。	
	□　疲労感はない	5
	□　少し疲労感がある	4
	□　ある程度疲労感がある	3
	□　かなり疲労感がある	2
	□　極めて疲労感がある	1
Q8	あなたの「ワークライフバランス」についてもっともあてはまるものを選んでください。	
	□　ワークライフバランスに満足している	5
	□　ワークライフバランスにやや満足している	4
	□　ワークライフバランスに満足しているかどうか，どちらとも言えない	3
	□　ワークライフバランスにあまり満足していない	2
	□　ワークライフバランスに満足していない	1
Q9	あなたの「キャリア」についてもっともあてはまるものを選んでください。	
	□　今のキャリアに満足している	5
	□　今のキャリアにやや満足している	4
	□　今のキャリアに満足しているかどうか，どちらとも言えない	3
	□　今のキャリアにあまり満足していない	2
	□　今のキャリアに満足していない	1

2） 妥当性

① 構成概念妥当性

Q1～Q9の回答結果とQOL関連指標（Q10～28）の相関係数を**表２**に示す。Q2の関連指標（Q13／Q14）との相関係数（0.157）以外は関連指標と一定の相関係数を示し，Q2についてはQ19／20やQ27とは一定の相関が認められたことから，Q1～Q9の構成概念妥当性は確認されたと判断した。

② 項目間の相関関係

各質問項目間の相関関係は0.3から0.7の範囲であった。

3） 信頼性

Q1～Q9の回答結果から算出されたCronbach α係数は0.897であった。各質問の必要度を確認するために，質問全体からQ1～Q9の１項目ずつ除いた場合のCronbach α係数を算出した結果，その変動は0.878から0.897の範囲であり，特段のα係数の変化は認められなかった。

4） 3rd質問票

定量調査の結果等について専門家と検討した結果，Q4のシーリング効果に対し４名の医師インタビューを実施し，回答形式が単極性であることが原因として推察されたため双極性に変更した。また，全体に時間的要素を限定した方がよいとの指摘から，「直近１カ月のこと」という期間を加え，さらに基準関連妥当性の確認のため（Q10）VASで仕事関連のQOLを測定することとした[8]。

（４） 確認調査

1） シーリング効果

Q4は，回答方式を双極性にしたことで回答が分散されシーリング効果は回避できたと考えられ，その他の質問においても回答の偏りは確認されなかった。

2） 基準関連妥当性

Q10（VAS）の回答は平均値が64.7，中央値が70.0，最小値が15，最大値が100であった。各質問項目とVAS間の相関係数は，Q8が0.793，Q6が0.700，Q5が0.675の順に高く，Q3が0.250，Q4が0.468，Q2が0.531の順に低かった。単純合計した全体の値とVASの平均値のリトワイズ法による相関係数は0.789であった。

3） 信頼性

Cronbach α係数は0.899であった。

4） 最終版質問票

3rd質問票におけるシーリング効果の回避，基準関連妥当性および信頼性の結果について，専門家との検討会を実施し最終化を図った。項目は，「仕事量」「仕事時間」「協力」「診療」「労働条件」「労働環境」「心理面」「WLB」「キャリア」の９項目で，回答形式は５段階から構成された質問票を最終版とした（**図２ａ**）。

Ⅳ 考　　察

医師の仕事関連のQOLを測定するツールの必要性を見いだし，オリジナルのプロファイル型尺度を４ステップで開発[4]し９項目５段階の質問票として最終化した。

第３ステップにおいて，Q2仕事時間の検証質問との相関係数が0.157と低かったのは，本来の仕事の問題認識が時間的尺度のみではとらえられないことが示唆された。各質問項目間および各質問項目と全体合計間の相関において内部妥当性は検証され，信頼性も担保されたと推察された。以上より修正を加え，基準関連妥当性を図るためにVASの質問を追加し3rd質問票としたが，VASで絶対的QOLを測定できたかは限界であった。

第４ステップで，シーリング効果の回避，Cronbach α係数は0.899で信頼性の担保は確認され，基準関連妥当性についても限界はあるがほぼ検証できたと考えてよいだろう。別途，スコア化について検討したところ，単純合計値とVASの相関係数が0.789，１因子分析との相関係数が0.782であることから，スコアを重みづけせず単純合計する（**図２ｂ**）ことについても問題はないと考えられた。

最後に，本研究は新規に医師仕事関連のQOLを測定するツールとして開発し，単純合計でスコア化して使用できるものとなった。予期せずコロナ禍での開発となりWebインタ

ビューとなったことやコロナによる変化や「医師の働き方改革」がうたわれる時期と重なったことなどの影響を加味していないことは限界の１つであろう。また，基準となる尺度がないことで絶対的な評価に限界があったことは弱点と考えており，今後は調査を重ねデータを蓄積して検討と改善をしていくことが必要である。

Ⅴ 結　　語

患者以外のQOL評価の必要性[1)]，医療資源のひっ迫[9)]，さらには「医師の働き方改革」関連法が2024年から施行されること[10)]を考え合わせれば，医療従事者の負担をコスト以外の観点から評価することの重要性が高まることは疑いない。医療従事者・供給者側のQOLは，これまでは評価軸からは除外されてきた。しかし，医療従事者側の負担をより高い解像度で可視化するのであれば主観的な評価測定のために，PROを拡張した概念として，PROVIDER Reported Outcomeの尺度を構築することが最適な解決法となろう。今回構築した尺度によって，これまで顧みられなかった医療従事者側のベネフィットを可視化することが可能になり，新たな価値評価軸の確立のためにも有意義な成果と考える。

謝辞

本研究の立ち上げにご協力いただきました株式会社インテージヘルスケア　マイケル・ロプレスティ氏，本研究分野における専門家としてアドバイスをいただきました岡山大学大学院保健学研究科教授齋藤信也先生（現同特命教授）および東北大学大学院医学系研究科准教授鈴鴨よしみ先生（現立命館大学）に厚く御礼申し上げます。

本研究は，大塚ホールディングス株式会社の資金を受け実施したものです。

滕麗達は大塚ホールディングス株式会社から研究費を受領し，五十嵐中は大塚ホールディングス株式会社，大鵬薬品工業株式会社から本研究以外で研究費を受領しています。

文　　献

1）中野陽介，廣實万里子，五十嵐中．医薬品の多様な価値－国民視点および医療環境変化を踏まえた考察－．医薬産業政策研究所リサーチペーパー・シリーズ　2022；No.79.

2）Ezura M, Sawada K, Takushima Y, et al. A systematic review of the characteristics of data assessment tools to measure medical doctors' work-related quality of life：Journal of Market Access & Health Policy 2023；11.

3）保健医療経済評価研究センター．QOL尺度（preference-based measure：PBM）．（https://c2h.niph.go.jp/tools/pbm/index.html）2023.11.4.

4）木原雅子，加治正行，木原正博．医学的測定尺度の理論と応用．東京：メディカル・サイエンス・インターナショナル，2020；6-194.

5）厚生労働省．平成30年医師・歯科医師・薬剤師統計の概況：結果の概要１医師．（https://www.mhlw.go.jp/toukei/saikin/hw/ishi/18/dl/kekka-1.pdf）2023.11.4.

6）Patrick DL, Burke LB, Gwaltney CJ, et al. Content validity--establishing and reporting the evidence in newly developed patient-reported outcomes（PRO）instruments for medical product evaluation：ISPOR PRO good research practices task force report：part 1--eliciting concepts for a new PRO instrument. Value in Health 2011；14.

7）COSMIN Risk of Bias checklist（July, 2018）.（https://www.cosmin.nl/wp-content/uploads/COSMIN-RoB-checklist-V2-0-v17_rev3.pdf）2023.11.4.

8）山田一成，江利川滋．Web調査におけるVisual Analogue Scaleの有効性評価．東洋大学社会学部紀要　2014；52：52-70.

9）厚生労働省新型コロナウイルス感染症対策推進本部．救急医療等のひっ迫回避に向けた対応について（事務連絡令和４年８月５日）．（https://www.mhlw.go.jp/content/000973982.pdf）2023.11.4.

10）厚生労働省．令和４年度第２回医療政策研修会第２回地域医療構想アドバイザー会議；資料14医師の働き方改革について．（https://www.mhlw.go.jp/content/10800000/001040964.pdf）2023.11.4.