

主論文の要旨

How many patients are required to provide a high level of reliability in the Japanese version of the CARE Measure? A secondary analysis

日本語版 CARE Measure において高い信頼性を得るためには
何人の患者からの評価が必要か？ 二次分析の結果

名古屋大学大学院医学系研究科 総合医学専攻
発育・加齢医学講座 総合診療医学分野
(指導：葛谷 雅文 教授)

松久 貴晴

【緒言】

一般診療において、医師の患者に対する共感は、診療効果を高める重要な要因である。医師の共感は患者の満足度を上昇させ、患者の服薬アドヒアランスを向上させ、患者自身の疾患の理解や治療への取り組みを強化させることが報告されている。また、医師の患者に対する共感のレベルが高いほど、患者の治療結果がよりよい傾向にあることも実証されている。

これまでに、医師の共感のレベルを評価するために多くの尺度が開発されてきた。しかし、そのほとんどは医師による自記式、もしくは観察者により評価されるものであるため、評価者の共感に対する認識のレベルが影響するという問題があった。診療は患者-医師の相互関係において成り立ち、医師の共感のレベルは患者により感じられるものであるため、医師の共感のレベルを評価するためには患者からの評価が最も妥当である。

The Consultation and Relational Empathy (CARE) Measure はプライマリ・ケア領域において、共感の受け手である患者が医療者の共感のレベルを評価するために作られた質問紙である。CARE Measure は英国で開発され、これまでに中国、オランダ、スウェーデン、クロアチアといった諸国で翻訳版が作成され、それぞれその妥当性と信頼性が検証された。日本でも名古屋大学医学部附属病院総合診療科で日本語版 CARE Measure が開発され、その妥当性と内的信頼性が検証された。しかし、英語版、中国版 CARE Measure と異なり、日本語版 CARE Measure は何人の患者が質問紙に答えることによってその医師の共感のレベルを高い信頼性を持って評価できるかについては検討されていない。そこで本研究では、日本語版 CARE Measure において、何人の患者から回答を得るとその医師の共感のレベルを高い信頼性を持って評価できるのかを明らかにすることを目的とした。

【方法】

先行研究で収集・集計された、日本語版 CARE Measure の患者回答データと患者背景調査データを用いて二次解析を行った。データ収集は 2011 年に名古屋大学医学部附属病院総合診療科で行った。同総合診療科の医師 9 人が研究に参加した。参加医師は、医師としての経験年数が 6 年目から 33 年目の医師であり、全員男性であった。9 人のうち、3 人は専攻医であり、2 人は医員であり、4 人は教員であった。対象患者は、2011 年の 7 月から 12 月まで、各々の医師を受診した成人患者である。担当医師が不安障害など患者の病状に影響を与えると判断した患者や、認知症など適切に回答できないと考えた患者は除外した。回答を得られた 317 名中、日本語版 CARE Measure の 10 項目すべての質問に回答した計 252 人を対象とし、解析を行った。

解析方法は、日本語版 CARE Measure の信頼性の指標として級内相関係数 (ICC) を評価した。ICC は下記の通り定義される。

$$ICC = \frac{\sigma_{GP}^2}{\sigma_{GP}^2 + \sigma_P^2}$$

σ_{GP}^2 は参加した総合診療医の日本語版 CARE Measure の平均点の分散、 σ_p^2 は患者間の分散である。患者のサンプルサイズが n の場合、 σ_p^2 は下記の通りとした。

$$\sigma_p^2 = \frac{\sigma^2}{n}$$

σ^2 は参加したすべての患者の日本語版 CARE Measure の総得点の分散である。このとき、一般化可能性理論における G 研究として、一元配置分散分析を行い、 σ_{GP}^2 と σ^2 を算出し、ICC（一般化可能性理論の G 係数に相当）を算出した。解析にはソフトウェアである G-string IVを用いた。今回 G 研究において、医師間の共感のレベルの差異を識別することを目標とし、評価者である各患者は 1 人の医師しか測定しないモデルとした。その後、一般化可能性理論の D 研究として、中国語版 CARE Measure で ICC を評価した先行研究に習い $ICC \geq 0.80$ を満たす患者数を求めた。

【結果】

Table 1 では各々の総合診療医を受診した患者の特性と日本語版 CARE Measure の得点を示す。研究に参加した患者数は、一人の医師に対して最少で 9 人、最多で 50 人であり、平均は 28 人であった。各総合診療医を評価した患者群ごとの特性では、年齢に有意差を認めた（ANOVA; $p < 0.0001$ ）。しかし、日本語版 CARE Measure の得点と患者の年齢に相関はなく、今回の解析結果に患者の年齢の影響は大きくないと考えられた（Spearman's rank coefficient = 0.001; $P = 0.990$ ）。医師毎の日本語版 CARE Measure の平均点は最小値で 34.8 点、最大値で 45.2 点であった。全患者における日本語版 CARE Measure の平均点は 38.8 点であった。

一元配置分散分析による推定値は、 $\sigma_{GP}^2 = 6.942$ 、 $\sigma^2 = 66.030$ であった。D 研究の結果、一人の総合診療医に対して、38 人以上の患者の評価で高い信頼性を持って共感のレベルを評価できると算出された（ $ICC \geq 0.8$ ）。（Table 2）

今回のデータは医師ごとに参加した患者数が大きく異なるため、20 人以上の患者に評価された医師のみを対象に解析を行った。38 人の患者の評価での ICC は 0.78 であり、 $ICC \geq 0.8$ は満たさないも許容内の結果と考えられ、38 人という結果は適切であると考えた。

また、すべての医師が 38 人の患者から評価されたと仮定した場合、医師間の平均点の平均は 38.6 点であり、標準偏差は 3.2 であった。四分位範囲は 36-42 点であり、42 点以上で有意に点数が高く、36 点以下で有意に点数が低いと考えられた（Figure 1）。

【考察】

日本語版 CARE Measure において、38 人以上の患者の評価で医師の共感のレベルを高い信頼性をもって評価ができることが明らかになった。38 人の患者から日本語版 CARE Measure を回収することは、日常臨床において、実現可能性の高いものである。

中国語版 CARE Measure では約 30 人の患者からの評価をもって医師の共感のレベルを測定でき、本研究結果と近い結果となった。英語版 CARE Measure では、本研究

や中国版と解析方法が異なり、参加医師ごとに ICC が 0.8 を超える患者数を算出している。英語版 CARE Measure の先行研究に参加した総合診療医のうち、必要とする患者数が最も多かった医師は 50 人であった。我々は同じ方法を使用して参加した医師毎の ICC を求めた。最も必要とする患者数が多かった医師は 53 人であり、結果は類似していた。

本研究の結果は、日常臨床で医師の共感を評価する有力な方法を示した。日本語版 CARE Measure を用いることで、日常臨床で患者の協力を得られれば、比較的簡易に医師の共感のレベルを高い信頼性をもって測定することができる。しかし、本研究にはいくつか研究の限界がある。はじめに、患者の選択にあたり、各医師を受診した患者を連続して選択しており、無作為に抽出しておらず、目的とする母集団を適切に代表するサンプルと言えない可能性がある。また、研究の実現可能性のため、特別な病气（不安障害や認知症）の患者は除外した。そして、本研究は単一施設（名大病院）で実施されており、地方の病院や開業医などでの評価は行っていない。大学病院での診療はその診療時間やケアの継続性から、日本のプライマリ・ケアの代表と言えない可能性がある。最後に、参加した医師が 9 人であり、これは先行研究と比較すると少なかった。

本研究より、日本において、態度評価の一つの指標としての共感の評価において、日本語版 CARE Measure が有力な方法として使用できることが立証された。日本版 CARE Measure を用いた更なる共感、そして患者中心の医療に対する研究の発展が期待される。

【結語】

本研究によって、日本語版 CARE Measure では一人の総合診療医あたり 38 人以上の患者の評価により医師の共感のレベルを高い信頼性をもって測定することができることが明らかになった。