

特集：「行動療法研究」における研究報告に関するガイドライン  
(展望)

## 尺度翻訳に関する基本指針

稲田 尚子<sup>1,2</sup>

### 要約

患者報告式アウトカム尺度が翻訳されることにより、研究成果の国際比較および国際共同研究が可能となる。翻訳された尺度から得られるデータの質は、その翻訳の正確さに依存する。ISPOR (International Society for Pharmacoeconomics and Outcomes Research) タスクフォースによるガイドラインは、患者報告式アウトカム尺度の翻訳に関する質を担保し、ひいては研究報告の質を高めるうえで有用となる。本稿では、当該ガイドラインに従って、推奨される10の手続き（(1) 事前準備、(2) 順翻訳、(3) 調整、(4) 逆翻訳、(5) 逆翻訳のレビュー、(6) 調和、(7) 認知デブリーフィング、(8) 認知デブリーフィング結果のレビューと翻訳終了、(9) 校正、(10) 最終報告）について解説した。また、当該ガイドラインに基づく具体的な記載事例の紹介を行い、尺度翻訳の際の留意事項について考察した。

キーワード：患者報告式アウトカム尺度 翻訳 研究報告の質 ISPORタスクフォース

### はじめに

研究成果の国際比較および国際共同研究を行う、あるいは、グローバルスタンダードである尺度を日本でも利用可能にする、という観点から、近年わが国では、海外、特に米英で開発された尺度の翻訳作業が相次いでいる。翻訳された尺度から得られるデータの質は、その翻訳の正確さに依存する。したがって、尺度翻訳の質を担保し、研究報告の質を高めるために、尺度翻訳の手続きに関する適切なガイドラインが必要である。本稿では、尺度を翻訳し、また必要に応じて自国の文化に適したものに変更する際に留意すべき事項を解説し、その記載事例を紹介するとともに、一連の翻訳プロセスに関する留意事項を考察する。具体的には、患者報告式アウトカム尺度の翻訳のガイドラインである、ISPOR (International Society for Pharmacoeconomics and Outcomes Research) タスク

フォースによる報告書 (Wild et al., 2005) に準拠し、尺度翻訳の際に留意すべき10の手続きに焦点を当てて解説する。当該ガイドラインは、本稿執筆時点 (2015年3月2日) で、Web of Scienceにおける被引用数が433編あり、すでに多くの研究者によって活用されている。なお、患者報告式アウトカム尺度とは、患者 (研究対象者) の健康状態に関して、専門家や他者の修正や下位尺度を介さない、患者自身の直接的な報告による測定のことである (U.S. Food and Drug Administration, 2009, p. 32)。つまり、当該ガイドラインは情報提供者が患者「本人」である尺度を対象として作成されたものであるが、情報提供者が患者の「親」や「教師」である尺度の場合にも、本稿で述べる尺度翻訳における基本指針は適用可能であると考えられる。

### ISPORタスクフォースによるガイドライン

尺度の翻訳に関しては、次のような四つの課題が指摘されていた (Wild et al., 2005)：(1) 翻訳の方法論に一貫性がない場合には、同じ尺度間の比較が困難になる、(2) 同じ手続きが異なる

<sup>1</sup> 東京大学大学院教育学研究科

<sup>2</sup> Southwest Autism Research & Resource Center  
(2015(平成27)年3月18日受理)

る用語で記述されている場合があると、内容の明確さに欠ける、(3) 尺度開発者が翻訳の独自ガイドラインを作成している際には、一貫性を欠き、時代にそぐわない場合がある、(4) 適切な翻訳プロセスを経ていない尺度は、収集されたデータの妥当性を脅かし、国際共同研究を行う際に不適切である。これらの理由から、尺度翻訳のベストプラクティスのために、ISPORタスクフォースによりガイドラインが作成された。ISPORタスクフォースは、先行研究における尺度翻訳の方法および生じている混乱の原因を調べ、ガイドライン作成のための方針として、以下の五つを定めた。

(1) **用語** 用語を整理する。先行研究では、例えば、翻訳後に5～6名に行う事前テストのことを、認知デブリーフィングやパイロット調査などと記述しており、用語に一貫性がない。通常、パイロット調査とは、大規模データを収集する前に行う30～40名を対象とした調査のことを指すため、翻訳手続きの記述には適切ではない。ガイドラインでは、用語を整理し、適切な用語を選択して記述することとした。

(2) **方法論** 方法論を整理する。先行研究では、例えば、逆翻訳（翻訳した尺度を別の翻訳者に元言語に翻訳してもらい、原著者が表現の等価性を確認すること）を行っている場合とそうでない場合があり、一貫性がない。逆翻訳は、誤訳の有無、表現の等価性を検討するために重要な手続きであるため、ガイドラインでは、翻訳に必要なプロセスとして含め、またその他の方法論も整理することとした。

(3) **翻訳対象** 異なる言語への翻訳に関する方法論について記載する。翻訳には、他言語への翻訳（例：英語から日本語へ）および、同言語のものを別の国で使用する際の翻訳（例：スペインで開発された尺度をアルゼンチンで使用するスペイン語版に翻訳する）がある。また、患者報告式の方法は多様であり（例：症状評価チェックリスト、日記、ケースレポート方式など）、報告方法の種類を特定する必要がある。

ガイドラインでは、他言語への翻訳方法に関して報告することとし、また、日記、ケースレポート等は、アプローチ方法の妥当性や厳密性に欠けるため、今回は症状評価尺度に焦点を当てることとした。

(4) **記述方法** 基準などを設けず、必要な手続きを記述する。ガイドラインの記載方法として、具体的な基準を設定するかどうか決める必要があったが、ほかの言語の性質や尺度の多様性を考慮して、ガイドラインでは、基準などを設けないこととした。それゆえ、ガイドラインは、標準的な尺度翻訳の方法について、誤った用語を用いることなく的確に記述され、ベストプラクティスが提示されているが、多少柔軟に用いることができる余地が残されている。

(5) **ガイドラインのユーザーによる翻訳プロセスの提示方法** 特定の尺度の翻訳を行った場合、最後にその一連の手続きについて報告する必要がある。その際の最終報告書の書き方には、厳密な基準を設けていないが、後述する原則に従い、明瞭に報告することとした。

以上を踏まえ、最終的に、尺度翻訳のベストプラクティスとして、以下の10の手続きが提案された（詳細は後述する）：(1) 事前準備、(2) 順翻訳、(3) 調整、(4) 逆翻訳、(5) 逆翻訳のレビュー、(6) 調和、(7) 認知デブリーフィング、(8) 認知デブリーフィング結果のレビューと翻訳終了、(9) 校正、(10) 最終報告 (Wild et al., 2005)。

### 尺度翻訳の原則

尺度翻訳に関する以下の10の手続きについて、内容、合理性、適任者、実施しないリスクの点からの原則を述べる。なお、Wildら (2005) の報告に準拠しつつ、海外で開発された尺度を日本語に翻訳する際の手続きとして解説する。

(1) **事前準備** 一連の尺度翻訳のプロセスの最初に行うべき重要な手続きである。プロジェクト責任者はまず原著者に日本語翻訳の許

可を求める。無許可で翻訳した場合は、著作権の侵害にあたり、その翻訳版も正式に承認されていないいわゆる海賊版となるため、確実にこの手続きを踏む。正式に翻訳の許諾を得た後には、当該尺度の概念や項目内容に関する誤解や曖昧さをなくすため、原著者らとやり取りして、尺度に対する理解を十分に深める。

(2) 順翻訳 原版の言語から日本語への翻訳を行う。母語が日本語で、原版の言語（通常では英語）に精通している翻訳者（日本在住で、尺度翻訳の経験がある人が望ましい）を2名たてる。翻訳者が1名であると、その翻訳者の独自のスタイルが多分に反映されてしまうため、2名が独立して翻訳することが推奨される。プロジェクト責任者は、翻訳者およびプロジェクト担当者に尺度が開発された背景や構成概念などを事前に説明し、尺度内容に対する誤解を防ぎ、共通理解を得るよう努める。また、順翻訳の際には、原版の意味を損なうことなく、日本語として自然であり、回答者が容易に理解できる表現を使うように十分に配慮する。また、原版の文化特有の表現や固有名詞が出てきた場合は、原著者と協議のうえ、日本文化に即したものにする。

(3) 調整 2名以上による順翻訳版を比較・統合し、一つの版を作成する。プロジェクト担当者および翻訳者間で議論を行い、誤訳を修正するほか、語の用い方など翻訳者の個人的スタイルに偏らないように調整する。順翻訳に関わっていない別の翻訳者がこの作業を行う場合もあれば、プロジェクト担当者が責任者と議論しながら行ってもよい。訳が難しい項目は別の翻訳者に確認する場合もある。

(4) 逆翻訳 順翻訳版を原版の言語に翻訳する。これは順翻訳された尺度の項目表現が原版と等価な概念・意味を持つ尺度であるかを原著者らに確認してもらうためである。順翻訳作業に携わっておらず、逆翻訳の経験がある翻訳者が適切である。

(5) 逆翻訳のレビュー 順翻訳の質を評価

するため、原版の著者らによって、逆翻訳されたものを原版と比較し、双方が等価であるかどうかレビューされる。乖離がある場合は、日本語翻訳あるいは逆翻訳のプロセスのいずれに問題があるのかを順翻訳者が検討し、必要な修正を行う。逆翻訳およびそのレビューを経していない翻訳版は、原版との等価性が不明であり、ひいては測定結果に大きな影響を与え、国際比較や国際共同研究も不可能となる。

(6) 調和 この手続きは、原版の著者らが行うものである。原版が複数の言語に翻訳されるプロジェクトが進行している場合は、原著者らは、複数の逆翻訳版と原版をそれぞれ比較し、項目表現が等価であるかどうかを一貫したやり方で一斉に検討する。これは異なる言語間で生じうる翻訳の問題を発見し修正するためである。これにより、原版と翻訳版の尺度の等価性を担保することができ、国際間におけるデータの蓄積、比較が可能となる。具体的には、原著者らは、各国の逆翻訳版を比較し、各項目表現が同等な意味を有しているか、および項目の概念に異なる言語間で乖離が生じていないかを確認する。それらが見つかった場合には、ほかの国のプロジェクト責任者に連絡し、該当項目について再度確認してもらうよう求める。これは後述する(8)の翻訳終了までのどの段階においても行われるが、主に逆翻訳のレビューの段階で行われる。同じ尺度の複数の言語への翻訳がほぼ同じタイミングで行われている場合は、この調和のプロセスを踏むことが比較的容易であるが、通常は各国で翻訳の時期がずれている場合が多い。そのような場合は、ほかの言語の翻訳時に生じた問題を記録しておき、次の他国の翻訳の際にも同じ問題が生じていないかを確認することになる。

(7) 認知デブリーフィング 使用が想定される日本語を母国語とする患者または一般母集団で少人数の調査を行う。これは、日本語翻訳の段階で気づかれなかった翻訳の問題を発見し、また実際のユーザーに対する日本語版の理

解しやすさ、認知的等価性を検討することが目的である。調査の対象者は、日本語を母語とし、尺度が想定する対象を代表する(例:診断、性、年齢、教育歴)5~8名とする。対象者に、実際に翻訳された尺度に回答してもらい、その後わかりにくい項目がなかったか、項目の内容や概念の理解は適切かを確認し、吟味する。これを実施しない場合には、回答できずに欠測項目が増えたり、項目の意味の誤解に基づいた回答が得られることになり、データの正確性に欠くことになる。

(8) 認知デブリーフィング結果のレビューと翻訳終了 プロジェクト責任者および担当者は、認知デブリーフィングの結果を総合して、必要に応じて項目表現を修正する。これにより、回答者が理解しやすく、日本語としてより自然な表現に改善することができる。表現の修正が生じた場合には、該当箇所について再度逆翻訳を行い、原著者らにレビューしてもらう。これらの作業が終了後、翻訳作業を終了させる。

(9) 校正 プロジェクト担当者および/あるいは校正担当者は、日本語版を最終的に見直し、誤字脱字、文法的な間違いなどを修正する。

(10) 最終報告 最後に、プロジェクト責任者および/あるいは担当者は、尺度翻訳のプロセスについての報告書をまとめる。この作業を行うことにより尺度翻訳の手続きおよび用語の選択理由を正確に記録しておくことができる。

#### ISPORタスクフォースに基づく記載事例

ISPORタスクフォースに基づく尺度翻訳の記載事例を紹介する。Nishiyama et al. (2014)は、自閉症スペクトラム障害 (Autism Spectrum Disorder) の症状の程度を量的に評価する複数の尺度を包括的に比較した研究の中で、「Social Responsiveness Scale—Adult Self Report (SRS2-AS)」(Constantino & Gruber, 2012) および「Subthreshold Autism Trait Questionnaire (SATQ)」(Kanne et al., 2012)

の二つの尺度の一連の翻訳過程を下記のように述べている。

『原版との概念的等価性を担保するために、われわれは、“患者報告式アウトカム尺度の翻訳のためのグッドプラクティスの原則”(Wild et al., 2005)に、準拠した。Western Psychological Services (WPS) および最終著者 (SK) から翻訳の許可を得た後、SRS2-ASおよびSATQは、英語が堪能な日本人研究者 (SS, MK) 2名によって独立して順翻訳された。SRS2-ASの英語から日本語への翻訳に際しては、親記入式のSRS成人用 (Constantino & Gruber, 2012)の日本語版がすでに開発されており、両尺度の比較可能性を確保するため、可能な範囲でSRS-A日本語版から訳文を採用するようにした。翻訳者らは、第一著者 (TN) とともに、二つの翻訳版を比較、調和させ、最終的に一つの翻訳版を作成した。その後、日本語が堪能なネイティブの英語翻訳者によって、逆翻訳の前あるいは途中で尺度の原版に触れることなく、英語に逆翻訳された。逆翻訳された尺度は、原版と日本語版の間に生じた訳の乖離を検討するために開発者(例:WPSあるいは最終著者)に送られた。われわれは、SRS2-ASに関しては5回、SATQに関しては4回この手続きを繰り返した。この手続きの中で、SATQでは翻訳が難しい項目が二つあった。まず、項目14 (social situations where I can meet new people and chat [e.g., parties, dances, sports, games])に関して、日本人のほとんどは日常生活であまりパーティを開かないという文化的差異が見られる。代わりに、コンパ (drinking get-togethers) や懇親会 (social gatherings) を行うため、項目14にはこれらの文言を加えた。さらに、英語でgameの言葉が意味するものを日本語で適切に言い換えるのは難しく、この文言を削除した。次に、項目15 (shrugging) に関して、日本人のほとんどは他者と話す際に肩をすくめる (shrug) ことがないため、この文言をどう扱うかについてのコンセンサスが得ら

れなかった。そのため、58名の健康な成人に対してフィールド調査を行い、項目15を含めたものと含めないものの2種類の翻訳版に回答してもらった。その結果、われわれの予想に反して、2種類の翻訳版の項目得点の分布は同等ではなかったため、この項目 (shrugging) はそのまま残すこととした。』

### 尺度の異文化適応に関する事例

先述したNishiyama et al. (2014) の記載事例からもわかるように、翻訳プロセスの中では、項目表現の等価性を保ちつつ、項目によっては日本文化に即した変更を行うことが特に重要である。尺度の内容に、欧米と日本との間で大きく異なる言語や文字体系が含まれている場合は、特に注意を払い、原著者らと十分に協議する。ここでは、これらの点に関して、筆者が開発に関わった尺度の事例を一部紹介する。なお、患者報告式アウトカムに限定せず、他者評価による尺度も含めている。

「Vineland-II 適応行動尺度 (Vineland Adaptive Behavior Scale-Second Edition: VABS-II)」(Sparrow et al., 2005) の項目において、日本文化に適したものに変更した例を一部紹介する。原版にある「不規則変化の複数形を使う (例: children, geese, mice, women)」は、日本語には複数形がないため「助数詞をほぼ正しく使う (例えば、「一頭」「二羽」「車が三台」など) に変更された。また、原版の「すべての活字体のアルファベットの大きい文字と小さい文字がわかる」は、日本版では「すべてのひらがながわかる」と変更され、識別できる文字の総数はほぼ同数となっている。また日本語のみに存在する撥音や促音に関して、「“がっこう”などの促音や“あそんだ”などの撥音が読める」という項目が日本語版には新たに追加された。「自閉症診断面接尺度改訂版 (Autism Diagnostic Interview-Revised: ADI-R)」(Rutter, Le Couteur, et al., 2003) および「対人コミュニケーション質問紙 (Social Communication

Questionnaire: SCQ)」(Rutter, Bailey, et al., 2003) では、人称代名詞の反転 (例えば、「私」と言う代わりに「あなた」や「彼・彼女」と言うなど) を評価する項目で大きな変更が行われた。日本語の日常会話場面では、主語をほとんど用いないため、これらの現象は見られないことが多い。協議の結果、原著者らの意見を踏まえ、日本語版では評価しないこととなった。具体的には、項目の訳文自体は残したまま、ADI-Rでは回答欄にはあらかじめ8 (非該当) が印字され、SCQでは回答欄が削除された。そのほか、SCQでは、子どもの対人的遊びへの参加の有無に関する項目で、遊びの例として“Here We Go Round the Mulberry Bush”と“London Bridge is Falling Down”が挙げられていた。後者は「ロンドン橋落ちた」という、すでに日本語に翻訳された歌があったが、前者の日本語の歌はなかったため、類似した内容で日本の子どもになじみのある遊びの「かごめかごめ」に置き換えた。これらの例からわかるように、原版の項目表現を日本文化に照らして等価なものに変更するだけでなく、原著者らとの協議を踏まえ、必要最低限の項目追加、削除を行う場合がある。なお、翻訳された日本語版は、その信頼性・妥当性を検討する必要があることは言うまでもない。

### 尺度翻訳に関する留意事項

筆者の経験や先行研究を基に、尺度翻訳に関する留意事項を述べる。

(1) 複数の翻訳版 原著者に翻訳の許可を求めることは翻訳プロセスの第一歩である。しかしながら、これまで翻訳された尺度の中には、この段階ですでに問題が散見され、結果として複数の翻訳版が存在する尺度がある。

第一に、原著者に翻訳の許可を得ているのか不明な尺度がある。例えば、自尊感情を測定する尺度として最もよく用いられている「Rosenberg 自尊感情尺度 (Rosenberg Self Esteem Scale: RSES)」(Rosenberg, 1965) は、

複数の日本語翻訳版が存在する（星野, 1970; Mimura & Griffiths, 2007; 山本ら, 1982など）。しかしながら、筆者の知る限り、発表された論文にはいずれも原著者から翻訳の許可を得たかどうかの記述がない。また、Mimura & Griffiths (2007) のみが逆翻訳を行っているが、原著者らのレビューは経っていないため、いずれの翻訳版が正式なものなのか判断できない。第二に、翻訳希望者が原著者から翻訳の許可を正式に得ているにもかかわらず、原著者が複数名に翻訳権を与えている場合がある。例えば、自閉症スペクトラム障害の症状の程度を量的に評価する「自閉症スペクトラム指数 (Autism-Spectrum Quotient: AQ)」(Baron-Cohen et al., 2001) の日本語版は2種類存在する。いずれも原著者から翻訳の許可を得て、順翻訳・逆翻訳・そのレビューという正式なプロセスを経て、さらにその信頼性と妥当性が報告されているが（栗田ら, 2004; Wakabayashi et al., 2006）、項目表現の違いは測定結果に影響を及ぼし、自閉症スペクトラム障害のスクリーニング目的のカットオフ値が異なるという事態が生じている。この場合、どちらの尺度がより適切あるいは原版と等価と判断すればよいのであろうか。また、複数の版の存在を知らない場合は、誤ったカットオフ値を利用する可能性があり、混乱をまねくことにもなる。尺度翻訳に際しては、プロジェクト担当者は先行研究のレビューを十分にを行い、原著者もまた開発した尺度の権利に関して十分に責任を持つ必要がある。

(2) 異文化間妥当性 理想的には、尺度翻訳が終了した後に、異文化間妥当性を検討することが望ましい。ここで、異文化間妥当性とは、「翻訳した尺度の項目の性能が、原版の尺度の項目の性能を十分に反映している程度」を意味する。異文化間妥当性の検討のためには、原版と類似の調査対象に回答を求め、二つの国で得られた回答を基に、確認的因子分析や項目反応理論を使用することで検討できる。詳細は本稿の範囲を超えるので、COSMIN (COnsensus-

based Standards for the selection of health Measurement Instruments) チェックリストを参照されたい (Mokkink et al., 2012)。

(3) 信頼性と妥当性の検討 先述した10の手続きで翻訳が完了した尺度は、当然ながらその日本語版を実際に日本の対象となる集団に適用した場合の精度（すなわち信頼性と妥当性の程度）を検証する必要がある。本特集の他稿（土屋, 2015）でその原則が述べられているが、尺度翻訳の権利を得た研究者は、同時に当該尺度の信頼性と妥当性を検討し、発表する責任を負うことを肝に銘じなければならない。筆者の知る限り、尺度の翻訳が10年以上前に完了しているにもかかわらず、その信頼性と妥当性の検討作業が遅れているために、未公表のまま実用化が待望されている尺度が複数ある。また、自閉症・発達障害児教育診断検査3訂版 (Psychoeducational Profile-Third Edition: PEP-3; Schopler et al., 2005) の日本版（茂木ら, 2007）は、出版されているものの、その尺度特性の検討および標準化作業は全く行われていない。現時点では、結果として求められる発達年齢は原版の米国のデータに基づく参考値にすぎないため、日本での標準化が待たれている。筆者にとっても耳が痛い話であるが、尺度を翻訳した研究者は、単に翻訳するだけではなく、必ずその信頼性・妥当性を検討する必要がある。

(4) 尺度の普及と改訂 尺度の翻訳、信頼性・妥当性の検証後は、当然ながらその尺度を臨床、研究の両場面に普及させるために、論文発表はもちろんのこと、そのほかにアウトリーチ活動や研修などを継続的に実施することが求められる。例えば、筆者が開発に携わった自閉症スペクトラム障害の早期スクリーニング尺度である「日本語版乳幼児期自閉症チェックリスト修正版 (Modified Checklist for Autism in Toddlers: M-CHAT)」(Inada et al., 2011) は、全国研修会やe-learningなどの提供が継続的に行われている（神尾, 2012）。その結果、現在

では日本の全国各地で、乳幼児健康診査の間診項目に導入され、自閉症スペクトラム障害の早期発見の一助となっている。尺度開発は、それ自体が「目的」ではなく、特定の症状を評価するための「手段」を手に入れることが目的であるため、尺度開発者は、発表および普及のための責任を負うことになり、安易な心構えで尺度の翻訳や開発に着手するべきではないだろう。

また、グローバルスタンダードとなった尺度は、研究知見の蓄積や時代の変化にあわせて改訂される場合がほとんどである。「Vineland-II 適応行動尺度」は、日本では2014年の10月によく刊行されたばかりだが（辻井ら, 2014）、間もなく米国で第3版が刊行されるため、その後すぐに日本版の開発が始まる予定である。日本版の開発が遅れると、「WPPSI (Wechsler Preschool and Primary Scale of Intelligence) 知能診断検査」のように、第2版が刊行されることなく、第3版の日本版の刊行が待たれているという事態が生じる。そして第3版の刊行までは、時代にそぐわないように感じられる検査を使用し続けなければならない。特定の尺度を翻訳、開発したからといって、その改訂版を翻訳、開発する義務は生じないが、ユーザーがそれを期待することは心に留めておく必要がある。

#### おわりに

本稿では、海外で開発された患者報告式アウトカム尺度を邦訳し、日本文化に即したものにするための基本指針を述べてきた。本稿により、ISPORタスクフォースによる報告の概要を理解した読者は、ぜひ原典に手を伸ばし、各項目の内容への理解を深めていただきたい。本稿が、臨床家と研究者にとって、尺度翻訳の手続きを振り返る機会となり、本邦からも良質なエビデンスが発信される一助となればと願っている。また、本稿の読者は、自分で尺度を翻訳する機会よりも、翻訳された尺度を実際に研究や臨床場面で使用する機会の方が多くと考え

られる。使用する尺度に関する先行研究をレビューする際に、尺度翻訳のプロセスの妥当性を判断するためにも本稿を役立てていただきたい。

#### 謝 辞

本稿の草稿に目を通し、詳細なコメントをくださった奥村泰之先生に深く感謝申し上げます。

#### 文 献

- Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. 2001 The Autism Spectrum Quotient (AQ): Evidence from Asperger Syndrome/high functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, 31, 5–17.
- Constantino, J. N. & Gruber, C. P. 2012 *Social Responsiveness Scale*. 2nd ed. Los Angeles, CA: Western Psychological Services.
- Goodman, R. 1997 The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, 38, 581–586.
- 星野 命 1970 感情の心理と教育 児童心理, 24, 1445–1477.
- Inada, N., Koyama, T., Inokuchi, E., Kuroda, M., & Kamio, Y. 2011 Reliability and validity of the Japanese version of the Modified Checklist for Autism in Toddlers (M-CHAT). *Research in Autism Spectrum Disorders*, 5, 330–336.
- 神尾陽子 2012 発達障害の子どもと家族への早期支援システムの社会実装：実装支援プロジェクト終了報告書 ([http://www.ristex.jp/examination/implementation/pdf/H24\\_kamio\\_houkokusho.pdf](http://www.ristex.jp/examination/implementation/pdf/H24_kamio_houkokusho.pdf)) (accessed March 9, 2015)
- Kanne, S. M., Wang, J., & Christ, S. E. 2012 The Sub-threshold Autism Trait Questionnaire (SATQ): Development of a brief self-report measure of sub-threshold autism traits. *Journal of Autism and Developmental Disorders*, 42, 769–780.
- 栗田 広・長田洋和・小山智典・金井智恵子・宮本有紀 2004 自閉症スペクトラム指数日本語版 (AQ-J) のアスペルガー障害に対するカットオフ 臨床精神医学, 33, 209–214.
- Mimura, C. & Griffiths, P. 2007 A Japanese version of

- the Rosenberg Self-Esteem Scale: Translation and equivalence assessment. *Journal of Psychosomatic Research*, **62**, 589–594.
- Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & de Vet, H. C. W. 2012 COSMIN checklist manual <<http://www.cosmin.nl/images/upload/File/COSMIN%20checklist%20manual%20v9.pdf>> (accessed March 9, 2015)
- Nishiyama, T., Suzuki, M., Adachi, K., Sumi, S., Okada, K., Kishino, H., Sakai, S., Kamio, Y., Kojima, M., Suzuki, S., & Kanne, S. M. 2014 Comprehensive comparison of self-administered questionnaires for measuring quantitative autistic traits in adults. *Journal of Autism and Developmental Disorders*, **44**, 993–1007.
- Rosenberg, M. 1965 *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- ラター, M., ベイリー, A., & ロード, C. 黒田美保・稲田尚子・内山登紀夫 (監訳) 2013 SCQ 日本語版マニュアル 金子書房 (Rutter, M., Bailey, A., & Lord, C. 2003 *The Social Communication Questionnaire*. Los Angeles, CA: Western Psychological Services)
- ラター, M., ル・クーター, A., & ロード, C. 土屋賢治・黒田美保・稲田尚子 (監修) 2012 ADI-R 日本語版マニュアル 金子書房 (Rutter, M., Le Couteur, A., & Lord, C. 2003 *Autism Diagnostic Interview-Revised*. Los Angeles, CA: Western Psychological Services)
- ショプラー, E., ランシング, M. D., ライクラー, R. J., & マーカス, L. M. 茂木敏夫 (訳) 2007 日本版 PEP-3—自閉症・発達障害児教育診断検査— (三訂版) 川島書店 (Schopler, E., Lansing, M. D., Reichler, R. J., & Marcus, L. M. 2005 *PEP-3: Psychoeducational Profile*. 3rd ed. Austin, TX: Pro-Ed)
- スバロー, S. S., チケッティ, D. V., & バッラ, D. A. 辻井正次・村上 隆 (監修), 黒田美保・伊藤大幸・萩原 拓・染木史緒 (監訳) 2014 Vineland-II 適応行動尺度 日本文化科学社 (Sparrow, S. S., Cicchetti, D. V., & Balla, D. A. 2005 *Vineland Adaptive Behavior Scales: Second Edition (Vineland-II)*. Livonia, MN: Pearson Assessments)
- 土屋政雄 2015 尺度研究の必須事項 行動療法研究, **41**, 107–116.
- U.S. Food and Drug Administration 2009 Guidance for industry: Patient-reported outcome measures: Use in medical product development to support labeling claims <<http://www.fda.gov/downloads/Drugs/Guidances/UCM193282.pdf>> (accessed March 9, 2015)
- Wakabayashi, A., Baron-Cohen, S., Wheelwright, S., & Tojo, Y. 2006 The Autism-Spectrum Quotient (AQ) in Japan: Cross-cultural comparison. *Journal of Autism and Developmental Disorders*, **36**, 263–270.
- Wild, D., Grove, A., Martin, M., Eremenco, S., McElroy, S., Verjee-Lorenz, A., Erikson, P., & ISPOR Task Force for Translation and Cultural Adaptation. 2005 Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: Report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value Health*, **8**, 94–104.
- 山本真理子・松井 豊・山成由紀子 1982 認知された自己の諸側面の構造 教育心理学研究, **30**, 64–68.

# Improving the Methodological Quality of the Translation and Cultural Adaptation Process for Patient-Reported Outcome Measures

Naoko INADA<sup>1,2</sup>

<sup>1</sup> Graduate School of Education, The University of Tokyo

<sup>2</sup> Southwest Autism Research & Resource Center

## Abstract

Translation and cultural adaptation for patient-reported outcome (PRO) measures allows a comparison of cross-cultural data and enables datasets to be pooled worldwide. The quality of the data derived from translated measures relies on the accuracy of the translation. A task force of the ISPOR (International Society for Pharmacoeconomics and Outcomes Research) provided useful guidelines for enhancing the quality of translation and cultural adaptation for patient-reported outcome measures. The present study aimed to enhance understanding of the translation and cultural adaptation process for patient-reported outcomes, based on the ISPOR task force guidelines, focusing on 10 steps: (a) preparation, (b) forward translation, (c) reconciliation, (d) back translation, (e) back translation review, (f) harmonization, (g) cognitive debriefing, (h) review of cognitive debriefing results and finalization, (i) proofreading, and (j) final report. In addition, a detailed explanation is provided based on published articles, and considerations relating to translation and cultural adaptation are discussed.

**Key Words:** patient-reported outcome (PRO) measures, translation, reporting quality, ISPOR task force