

Aus der
Klinik für Anästhesiologie mit Schwerpunkt operative Intensivmedizin
Campus Charité Mitte und Campus Virchow-Klinikum
der Medizinischen Fakultät Charité - Universitätsmedizin Berlin

DISSERTATION

Richtlinienkonforme Übersetzung der STARD-Checkliste
zur präzisen Dokumentation von Diagnosestudien

zur Erlangung des akademischen Grades
Doctor medicinae (Dr. med.)

vorgelegt der Medizinischen Fakultät
Charité – Universitätsmedizin Berlin

von

Michael Barz

aus Berlin

Gutachter/in: 1. Prof. Dr. med. C. Spies
2. Prof. Dr. med. Chr. Lehmann
3. Prof. Dr. med. R. Rossaint

Datum der Promotion: 07.09.2012

Vorveröffentlichung

Ergebnisse dieser Arbeit wurden mit Genehmigung der Medizinischen Fakultät, vertreten durch den Mentor der Arbeit, zur Publikation eingereicht:

J.-P. Breuer, M. Seeling, M. Barz, T. Baldini, K. Scholtz, C. Spies. STARD-Kriterien - Checkliste zum Berichten von Diagnosestudien - Eine Standardisierte Übersetzung der Standards for Reporting of Diagnostic Accuracy (STARD)* auf Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen (ZEFQ)

**Patrick M Bossuyt, Johannes B Reitsma, David E Bruns, Constantine A Gatsonis, Paul P Glasziou, Les M Irwig, Jeroen G Lijmer, David Moher, Drummond Rennie, Henrica C W de Vet for the STARD steering group. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative 'BMJ. 2003 Jan 4;326(7379):41-4.*

Inhaltsverzeichnis

Inhaltsverzeichnis	4
1. Einleitung	5
2. Methodik	9
3. Ergebnisse	23
4. Diskussion	30
5. Schlussfolgerung	39
6. Zusammenfassung	40
7. Literaturverzeichnis	41
8. Lebenslauf	46
9. Danksagung	48
10. An Eides Statt	49

1. Einleitung

Schlecht durchgeführte Studien, aber auch eine schlechte Beschreibung hinsichtlich Fragestellung, Durchführung, statistischer Analyse und Schlussfolgerungen einer Studie können zu völlig unterschiedlichen Interpretationen und Schlussfolgerungen beim Leser und in der Konsequenz schlimmstenfalls therapeutischen Entscheidungen führen, die die Gesundheit von Patienten gefährden. Eine exakte Beschreibung der Studie kann dazu beitragen schon im Vorfeld Missverständnisse, Zweideutigkeiten und damit Möglichkeiten der Fehlinterpretation zu minimieren und somit sogar Patienten schützen und das Gesundheitssystem vor zusätzlichen Kosten, wie zusätzlichen diagnostischen Maßnahmen aufgrund zweideutiger Testergebnisse, bewahren [9].

Für verschiedene Studientypen sind bereits Leitlinien in Form von Fragekatalogen entwickelt worden, deren Anwendung einen Mindeststandard beim Berichten über diese Studienformen sichern soll. Die meisten dieser Checklisten, wie zum Beispiel das CONSORT-Statement [1] zur Qualitätsverbesserung von Reports randomisierter Studien im Paralleldesign, das STROBE-Statement [2;3] zur Verbesserung der Berichte über nicht randomisierte Querschnitts-, Fall-Kontroll- und Kohortenstudien, das QUOROM-Statement [4;5] zur systematischen Übersicht randomisierter Kontrollstudien oder das MOOSE-Statement [6] zu Beobachtungsstudien wurden im englischsprachigen Raum verfasst und liegen zunächst nur im englischen Original vor.

Eine Übersetzung in andere Sprachen könnte die internationale Verbreitung solcher Checklisten fördern und die Qualität wissenschaftlicher Publikationen steigern.

Die „Translation and Cultural Adaptation group“

Allerdings birgt dieses die Gefahr von inhaltlichen Übertragungsfehlern, weshalb eine Übersetzung standardisierten Regeln folgen sollte. Die „International Society for Pharmacoeconomics and Outcome Research“ (ISPOR) hat sich mit dieser Problematik befasst und zum Zwecke einer sicheren Adaptation in andere Sprachen Richtlinien vorgeschlagen [7;8]. Die ISPOR wurde im Jahre 1995 gegründet. Eines der Ziele der Organisation ist die Verbesserung des Wissens- und Informationsaustauschs zwischen

verschiedenen im Gesundheitswesen tätigen Gruppen, wie Forschern, medizinischem Fachpersonal, den Medien und der Gesellschaft.

Innerhalb der ISPOR gibt es verschiedene Arbeitsgruppen, die sich mit unterschiedlichen Themen befassen. Im Jahre 1999 gründete die ISPOR die „Quality of Life Special Interest Group“ (QoL-SIG) – „Translation and Cultural Adaptation group“ (TCA-group) um eine Debatte über die Schaffung von Richtlinien und Standards für die Übersetzung von Ergebnissen medizinischer Forschung in eine andere Sprache unter Berücksichtigung kultureller Besonderheiten des anderen Sprach- und Kulturkreises anzustoßen. Übersetzungen medizinischer Veröffentlichungen hatten bis dahin keine standardisierte Grundlage.

Auf dem ersten Treffen der TCA-Gruppe in Antwerpen im Jahre 1999 kamen Vertreter der Pharmaindustrie, der Hochschulen sowie Auftragsforschungsinstituten (CROs) zusammen. Die Vorträge und Diskussionen auf diesem Treffen zeigten, dass viele verschiedene Vorgehensweisen zum Übertragen von wissenschaftlichen, medizinischen Texten in andere Sprachen benutzt werden, es aber keine allgemeinverbindlichen Regeln hierzu gibt. Die TCA-Gruppe beschloss, dass es sinnvoll sei bessere Regeln und Definition aufzustellen da:

1. in der Praxis verschiedene Methoden verwandt werden, um ähnliche Aufgaben zu erfüllen, so dass die Ergebnisse schwer vergleichbar und nicht konsistent sind.
2. unterschiedliche Terminologien für den gleichen Sachverhalt benutzt werden, was die Verständlichkeit erschwert.
3. Entwickler von Frageinstrumenten bisweilen eigene Richtlinien für Übersetzungen entworfen haben, welche veraltet sind oder den Ansprüchen gegenwärtiger Forschung nicht mehr genügen.
4. schlecht übersetzte Texte die Validität von Forschungsergebnissen kompromittieren.

Gleichzeitig wurden folgende Mängel und Diskussionsbedarf zu folgenden Punkten ausgemacht:

1. Mangel an Konsistenz in der Terminologie. So werden die Begriffe „Pilottestung“ und „kognitives Debriefing“ beide benutzt, wenn ein neu übersetztes Diagnoseinstrument an einer kleinen Gruppe von 5 oder 6 Patienten erprobt wird. Die „Pilottestung“ beschreibt aber auch erste Tests an kleinen Personengruppen von 30-40 Patienten.

2. Mangel an Konsistenz bei der Methodik. Einigkeit besteht darüber, dass eine Übersetzung inhaltlich möglichst deckungsgleich und an den jeweiligen Kulturkreis adaptiert sein sollte. Ob jedoch eine Rückwärtsübersetzung erfolgen sollte, um die Übersetzungsgenauigkeit zu verfeinern wird unterschiedlich beurteilt. Auch ist nicht festgelegt, ob eine Übersetzung von nur einer Person gefertigt sein sollte oder ob mehrere Fassungen abgeglichen werden sollten.
3. Der Geltungsbereich zukünftiger Leitlinien - sollte man insbesondere für die Übersetzung bestimmter PRO Instrumente (Symptomchecklisten, Teile von case report forms) sprachspezifische Methodiken oder Übersetzungsmethoden vorschreiben?
4. Sollten die Leitlinien empfehlenden Charakter haben oder Mindeststandards sein, die erfüllt werden müssen?
5. Definition der Zielgruppe für die Leitlinien

Während der Leitlinienentwicklung wurde beschlossen, dass Übersetzungsprozesse in eine andere Sprache (zum Beispiel englisch – deutsch, französisch – russisch, etc.) und sprachliche Adaptationen an einen Kulturkreis gleicher Sprache (Übertragung eines spanischsprachigen Dokuments für Spanier in Spanisch für Argentinier) getrennt betrachtet werden sollten.

Aufgrund der unterschiedlichen Zielgruppen wurde beschlossen, dass die neue Leitlinie nicht überreglementiert sein sollte. Man plante die Entwicklung einer „good practice Leitlinie“, welche einen Mindeststandard fordert, aber Flexibilität erlaubt. Sie sollte einem Punkteplan entsprechen, welcher Schritt für Schritt abgearbeitet werden kann und somit für jeden verständlich ist. In einem weiteren Teil sollte beschrieben werden, welche Gefahren das Auslassen einzelner Punkte des Übersetzungsprozesses birgt.

Die TCA-Gruppe begutachtete hierzu 12 bestehende Leitlinien zur Übersetzung und kulturellen Anpassung folgender Gesellschaften:

1. American Association of Orthopaedic Surgeons (AAOS) [10]
2. Association of Test Publishers [11]
3. EORTC group [12]
4. EuroQoL group [unveröffentlicht]
5. Evidence: Clinical and Pharmaceutical Research [13]

6. FACIT group [14]
7. Health Outcomes group (HOG) [15]
8. Health Utilities Inc. (HUInc) [16]
9. International Quality of Life Assessment (IQOLA) group [17]
10. Kidney Disease Quality of Life (KDQOL) [18]
11. Medical Outcomes Trust (MOT) [19]
12. World Health Organisation [20]

Alle Leitlinien wurden hinsichtlich verschiedener Schritte des Übersetzungsprozesses beurteilt (z. B.: Vorwärtsübersetzung, Konsensusübersetzung, Rückübersetzung, Pilottestung). Die TCA-Gruppe entwarf darauf basierend ein Arbeitspapier, welches nach mehrfacher Überarbeitung als Leitlinie für Übersetzungen „The Translation and Cultural Adaptation – Principals of Good Practice“ verabschiedet wurde.

Die ‘STARD-Initiative’

Im Jahre 2003 entwickelten Bossuyt et al. [9] die „STAndards for Reporting of Diagnostic accuracy“ (STARD), einen 25 Punkte umfassenden Kriterienkatalog, um einen Mindeststandard für das Berichten über Diagnosestudien zu etablieren. Zweck dieser Mindeststandards ist für den Leser eine bessere Transparenz in Bezug auf Design, Durchführung, Analyse und Schlussfolgerungen einer Studie zu erzielen und diese damit sowohl nachvollziehbar zu machen als auch etwaigen Unklarheiten vorzubeugen. Allerdings lag auch dieser Katalog nur im englischen Original vor.

Ziel der Arbeit

Ziel war es, eine durch die englischsprachigen Originalautoren autorisierbaren Deutschen Version der STARD-Empfehlungen durch Anwendung der ISPOR-Kriterien zur standardisierten Übersetzung medizinischer Schriftinstrumente zu erstellen.

2. Methodik

Die Übersetzung der STARD-Kriterien folgte wie erwähnt den Empfehlungen der TCA-Gruppe der ISPOR. Diese sieht im Einzelnen folgende Schritte vor:

Tabelle 1 Arbeitsschritte gemäß der ISPOR-Empfehlungen

Arbeitsschritte	Inhalt
1 Vorbereitung (<i>Preparation</i>)	Kontaktaufnahme und Rücksprache mit den Originalautoren/Entwicklern mit Einholen der Einverständniserklärung zur Übersetzung des Originals
2 Vorwärtsübersetzung (<i>Forward Translation</i>)	Mindestens zwei von einander unabhängig zu erstellende Vorwärtsübersetzungen des Originals
3 Konsensusfindung (<i>Reconciliation</i>)	Gemeinsame Konsensusfindung über eine einzige vorläufige Version der Vorwärtsübersetzung aus allen einzelnen Vorwärtsübersetzungen
4 Rückübersetzung (<i>Back Translation</i>)	Rückübersetzung der vorläufigen Vorwärtsübersetzung in die Sprache des Originals durch einen staatlich geprüften Übersetzer/Dolmetscher, ohne Kenntnis der Originalversion
5 Review der Rückübersetzung (<i>Back Translation Review</i>)	Überprüfung der Rückübersetzung im Abgleich mit dem Original im Hinblick auf die konzeptionelle Übereinstimmung in Kooperation mit den Originalautoren/Entwicklern
6 Harmonisierung (<i>Harmonization</i>)	Bewertung der rückübersetzten Version zusammen mit den Originalautoren/Entwicklern und Konsensusfindung über konzeptionell strittige oder unklare Inhalte und Festlegung auf entsprechend nötige Revisionen
7 Kognitives Debriefing (<i>Cognitive Debriefing</i>)	Überprüfung der Verständlichkeit des Zielinstrumentes durch unabhängiges Fachpersonal (5–8 Personen)

8	Review des kognitiven Debriefings und Finalisierung (<i>Review of Cognitive Debriefing Results and Finalization</i>)	Umsetzung der Ergebnisse aus dem Debriefingprozess in eine erste finale Version
9	Korrekturlesen (<i>Proofreading</i>)	Überprüfung der finalen Fassung auf letzte Form-, Orthografie- und Grammatikfehler
10	Abschlussbericht (<i>Final Report</i>)	Zusammentragen der Resultate und abschließende Beurteilung, Beschreibung des Übersetzungsprozesses und Begründung aller darin getroffenen Entscheidungen

Modifiziert nach [12]; ISPOR = International Society for Pharmacoeconomics and Outcome Research

1. Vorbereitung

Hierbei handelt es sich um die Arbeit vor der eigentlichen Übersetzung, wie dem Einholen der Erlaubnis zur Übersetzung des Originaltextes oder der Ernennung von Verantwortlichen für die Übersetzung.

Zunächst wurde mit den Autoren der STARD-Kriterien Kontakt aufgenommen und um die Erlaubnis gebeten, eine deutsche Übersetzung der STARD-Kriterien zu erstellen. Diese wurde von Prof. Dr. Bossuyt als Repräsentant der STARD-Gruppe erteilt. Gleichzeitig wurden alle Originalautoren zur Mitarbeit an der deutschen Übersetzung der STARD-Kriterien eingeladen.

Die diese Dissertation betreuende Hochschullehrerin (Prof. Spies) bestimmte eine Gruppe von sogenannte ‚key in-country‘- Repräsentanten. Diese umfasste neben dem Promovenden drei weitere ärztlich-wissenschaftliche Mitarbeiter der eigenen Klinik. Diese Gruppe wurde mit der Aufgabe der Übersetzung betraut; allesamt sind deutsche Muttersprachler.

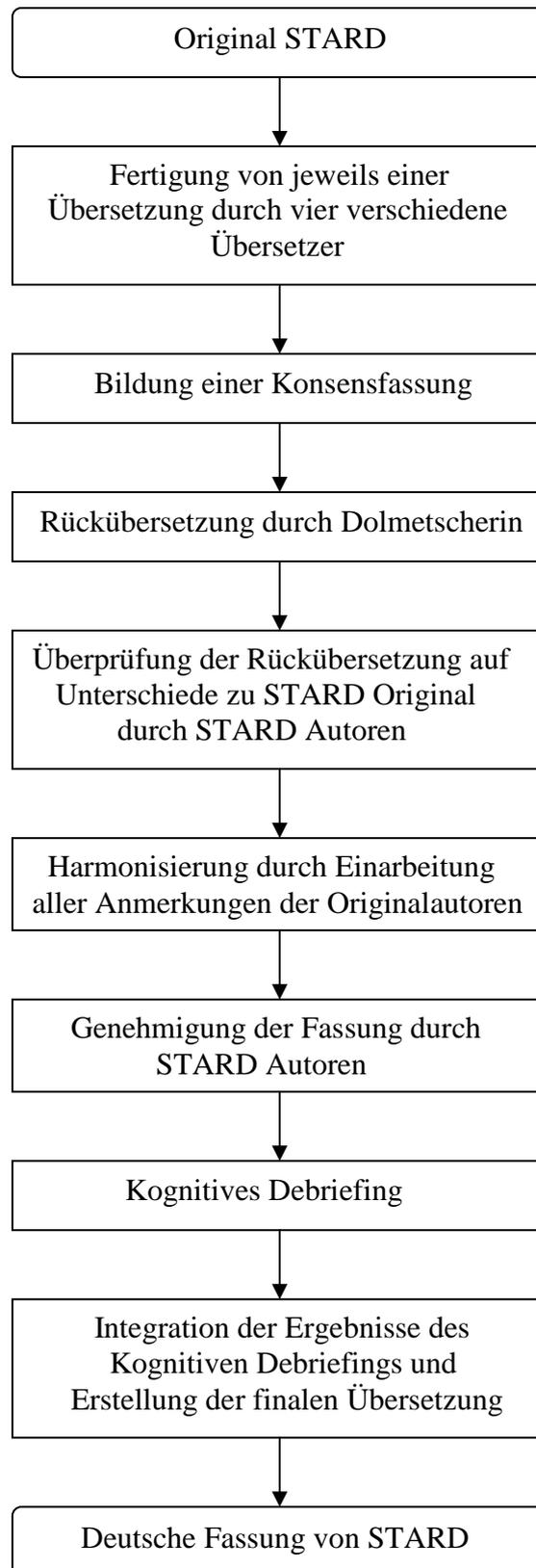


Abbildung 1: Flowchart des Übersetzungsprozesses

2. Vorwärtsübersetzung

Vorwärtsübersetzungen sollten von mindestens 2 Muttersprachlern (der in die zu übersetzende Sprache) unabhängig voneinander vorgenommen werden. Die Fertigung von nur einer Übersetzung birgt die Gefahr eines zu persönlichen, sprachlichen Stils des Übersetzers oder auch die systematische Falschübersetzung sich wiederholender Inhalte.

Der Promovend und die weiteren drei ‚key in-country‘- Repräsentanten übersetzten unabhängig voneinander die 25 Punkte der STARD-Checkliste. Dabei zeigten sich folgende diskussionswürdige Aspekte: Sollte die deutsche Version wörtlich möglichst nahe an der englischen Originalfassung bleiben? Solche Fassung klang eher hölzern anglistisch und entsprach nicht dem deutschen Sprachempfinden. Sollte die Fassung eher sinngemäß sein? Eine solche Fassung las sich besser, es stellte sich jedoch die Frage, ob diese Version dann nicht zu frei sei und diese dann noch einer eins- zu- eins- Übertragung entspräche.

3. Konsensusfindung

Die Konsensusfindung dient dem Ausräumen von Unterschieden bei den einzelnen Übersetzungen. Wie der Abgleich zu erfolgen hat, wird von der TCA-Gruppe nicht explizit vorgeschrieben. Individuelle Sprachgewohnheiten der Einzelübersetzungen sollen so nivelliert und eine sprachliche Konsistenz gewährleistet werden.

Im Rahmen der Konsensusfindung wurde aus den verschiedenen Übersetzungen eine gemeinsame Übersetzung gefertigt. An den insgesamt 4 Konsensustreffen nahmen stets alle Personen teil und nach Erörterung der strittigen Punkte fiel die Entscheidung auf eine Übersetzung einstimmig. Nicht festgelegt worden war, ob die Übersetzung wörtlich oder sinngemäß zu erfolgen hätte. Die ISPOR schlägt hierzu vor, dass eine Übersetzung möglichst für eine breite Schicht der Bevölkerung verständlich sein sollte, aber auch den Sprachgebrauch der Zielgruppe berücksichtigen sollte. Im gegebenen Fall bestand Einigkeit, dass man möglichst wörtlich am Original bleiben wollte. Zugleich sollte die Übersetzung aber auch sinngemäß sein, womit im Zweifel vom genauen Wortlaut abgewichen werden sollte, wenn dadurch der typisch deutsche Sprachgebrauch gewährleistet würde. Eine allgemeinverständliche Übersetzung wurde zwar angestrebt, da das Zielpublikum aus der medizinischen Fachwelt stammt, war die Verwendung von Fachsprache unumgänglich.

Weiterhin wurde beschlossen, den Begriff „Index- Test“ unverändert zu belassen. Er beschreibt das zu untersuchende, neue/weiterentwickelte Verfahren. Der Begriff „referencestandard“ wurde übersetzt mit Referenzverfahren. Dieses beschreibt das Verfahren gegen das der Indextest in seiner Genauigkeit das Vorliegen einer bestimmten Eigenschaft zu

erkennen getestet wird. Oftmals wird das Referenzverfahren der Goldstandard sein. Da aber denkbar ist, dass ein neuer Test nur gegen ein anderes Verfahren getestet wird oder aber Uneinigkeit darüber besteht, welches Referenzverfahren der Goldstandard ist, wählten wir den Begriff Referenzverfahren und nicht Referenzstandard.

4. Rückübersetzung

Um sicherzustellen, dass die Vorwärtsübersetzung inhaltlich dieselbe Bedeutung hat, wie die Originalfassung, soll eine Rückführung der konsentierten Vorwärtsübersetzung durch eine professionelle Übersetzung zurück in die Originalsprache erfolgen. Ob eine oder mehrere Rückübersetzungen gefertigt werden sollen wird von der ISPOR nicht festgeschrieben. Die Fertigung einer solchen Rückübersetzung soll eventuelle Unterschiede zwischen dem Original und der Vorwärtsübersetzung im Hinblick auf Inhalt und Bedeutung aufzeigen.

Die Rückübersetzung der STARD-Checkliste wurde durch eine staatlich geprüfte Übersetzerin und Dolmetscherin der englischen Sprache vorgenommen. Der Übersetzerin stand hierzu das englische Original nicht zur Verfügung.

5. Review der Rückübersetzung

Nach Meinung der TCA-Gruppe ist dies einer der wichtigsten Punkte, welchem aber in bisherigen Übersetzungsleitlinien kaum Beachtung geschenkt wurde. Die Rückübersetzung sollte mit dem Original verglichen werden, um Unterschiede in der Sprache, aber auch der Bedeutung aufzudecken und auszuräumen. Ziel ist das Erreichen einer gleichbedeutenden Übersetzung. Dieser Abgleich sollte durch den Projektmanager, seinen Mitarbeitern und wenn möglich mit den Originalautoren erfolgen und unklare Punkte beseitigen. Bei Nichtdurchführen dieses Schrittes droht das Verbleiben von Übersetzungsfehlern in der Zielsprache.

Die englische Rückübersetzung wurde im Rahmen eines Konsensustreffens von den vier deutschen ‚key in-country‘- Repräsentanten gegengelesen. Hierbei wurde die zurück übersetzte englische Version mit der englischen Originalversion hinsichtlich etwaiger Auffälligkeiten verglichen.

6. Harmonisierung

Sinn der Harmonisierung ist die Bewertung der zurück übersetzten Version zusammen mit den Originalautoren/Entwicklern und die Konsensfindung über konzeptionell strittige oder unklare Inhalte, und die Festlegung auf entsprechend nötige Revisionen.

Unter Leitung des Projektmanagers sollten alle Versionen bezüglich Übereinstimmung in Sprache und Bedeutung überprüft werden. Schwer zu lösende Probleme hinsichtlich der

Bedeutung sollten mit den Originalautoren erörtert werden. Unterschiedliche Fassungen bei Übersetzungen erschweren die Vergleichbarkeit von internationalen Datensätzen.

Die zurück übersetzte englische Fassung wurde mit der Original- und der deutschen Version verglichen. Die entstandene englische Übersetzung wurde anschließend den Originalautoren vorgelegt. Diese überprüften, ob die von der deutschen Fassung gefertigte englische Übersetzung sinngemäß der englischen Originalfassung entspricht.

Diese bis dato erfolgten einzelnen Schritte des Harmonisierungsprozesses wurden in standardisiert-strukturierter Form dokumentiert und dem Originalautor somit zusammengefasst zur Beurteilung vorgelegt: (a) Wiederholung seiner Kommentare, (b) die jeweils zugeordnete Fehlerkategorie, (c) die Kommentare und Argumentationen der Deutschen Autoren zur Harmonisierung sowie (d) deren Lösungsvorschläge und (e) die Beurteilung und jeweils abschließende Revision durch die Übersetzerin.

Die ‚key in-country‘- Repräsentanten erörterten auf zwei Treffen diskrepante Punkte und berieten sich hierüber weiter mit den Originalautoren.

STARD – Harmonization of the Certified translation from German to English (backward translation)

Issue	Description plus comments of the original authors	Discrepancies A = technical B = content	Comments of the german authors about the harmonization process	Proposal of solution	Translator's response and her revised backward translation	
1	Identify your article as survey of diagnostic accuracy (recommend the MeSH* terms "sensitivity and specificity" as key words).	A	Original (O): <i>study</i> , Forward Translation (FT): <i>Studie</i> , Backward Translation (BT): <i>survey</i> Die Übersetzerin hat einen spezifischen Term gewählt dessen allgemeine Zuordnung exakt dem Originaltext entsprechen hätte, d.h. jeder „survey“ ist eine studie aber nicht jede Studie ist ein „survey“. Translator chose a term, which is too specific for the present purpose, i.e. every "survey" is a "Studie" (study). However, not every "Studie" (study) is a "survey".	Frage an Übersetzerin (FaÜ): Könnte in diesem Zusammenhang sprachlich korrekt „Studie“ mit „study“ übersetzt werden? Question to translator (QtT): Would „study“ be a correct alternative translation to "survey" in this circumstance?	Ja /Yes Revision: Identify your article as a <i>study</i> of diagnostic accuracy (recommend the MeSH* terms "sensitivity and specificity" as key words).	Kommentar: Not identical to study, but close
2	Specify scientific problems [...] different survey groups.	Non	-	-	-	
3	Describe the survey population: inclusion and exclusion criteria as well as general conditions and places of data collection.	A	O: <i>setting</i> , FT: <i>Rahmenbedingung</i> , BT: <i>general conditions</i>	FaÜ: Könnte rein sprachlich in diesem Zusammenhang „Rahmenbedingung“ auch mit „setting“ übersetzt werden? Wenn nicht, dann Übernahme von „setting“ als (mittlerweile recht üblichen) Anglizismus in die deutsche Version. QtT: Could Rahmenbedingung be translated with setting in this circumstance?	Setting kann auch soviel wie Kulisse, Hintergrund, Szenario bedeuten (Film- und Theatersprache). Wenn das – im übertragenen Sinne – gemeint ist: ja/yes Revision: Describe the <i>study</i> population: inclusion and exclusion criteria as well as <i>setting</i> and places of data collection.	Kommentar: See above Kommentar: Again, sufficiently close

4	Describe the selection of test participants: Was the selection based on existing symptoms, on results of earlier tests or on the fact that test participants have undergone index tests or the reference methods? Note: Other selection methods are possible.	A	O: <i>presenting symptoms</i> , FT: <i>vorhandene Symptome</i> , BT: <i>existing symptoms</i>	FaÜ: Könnte in diesem Zusammenhang „vorhandene Symptome“ auch freier mit „presenting symptoms“ übersetzt werden? Qt: Would „presenting symptoms“ be a correct alternative translation to „vorhandene Symptome“ in this circumstance?	Ja /Yes Revision: Describe the selection of test participants: Was the selection based on <i>presenting symptoms</i> , on results of earlier tests or on the fact that test participants have undergone index tests or the reference methods? Note: Other selection methods are possible.	Kommentar: The English states „presenting“ – the symptoms with which patients present – no major problem
5	Describe the sample selection method for the participants: did the survey population correspond to an ongoing acceptance of participants as defined in selection criterion 3 and 4? If not, specify how the participants have been selected.	B	O: <i>consecutive series</i> , FT: <i>fortlaufende Aufnahme</i> , BT: <i>ongoing acceptance</i>	Änderung der dt. Version in: "... einem konsekutiven Einschluss von Teilnehmern ...?"; plus FaÜ: Kann „konsekutiver Einschluss“ mit „consecutive series“ in diesem Kontext übersetzbar wäre. Revision of german version into "... einem konsekutiven Einschluss von Teilnehmern ...?"; plus Qt: Can „konsekutiver Einschluss“ be translated with „consecutive series“ in this context?	Einschluss im Sinne von Aufnahme /Einbeziehen /Dazugehörigkeit ist nicht dasselbe wie „series“ (Reihe/Aufeinanderfolge) Nach Rücksprache: Ja, „consecutive series“ kann hier mit konsekutiver Einschluss übersetzt werden Yes Revision: Describe the sample selection method for the participants: did the survey population correspond to a <i>consecutive series</i> of participants as defined in selection criterion 3 and 4? If not, specify how the participants have been selected.	Kommentar: Is this sufficiently close to „consecutive series“?
6	Describe the data collection: [...] reference method?	Non	-	-	-	
7	Describe the reference [...] principle.	Non	-	-	-	
8	Describe the technical [...] methods.	Non	-	-	-	
9	Give the definition for [...] and reference methods.	Non	-	-	-	
10	Describe the number, [...] and the reference method.	Non	-	-	-	
11	Describe if the evaluators [...] available to the evaluator.	Non	-	-	-	
12	Describe the methods [...] (e. g.: 95% confidence interval).	Non	-	-	-	
13	Describe the methods [...] if accomplished.	Non	-	-	-	
14	Specify when the survey has been performed, including beginning and end of the selection.	A	O: <i>recruitment</i> , FT: <i>Rekrutierung</i> , BT <i>selection</i>	FaÜ: Kann „Rekrutierung“ auch mit „recruitment“ übersetzt werden? Qt: Can „Rekrutierung“ also be translated with „recruitment“, instead of „selection“ in this circumstance?	Ja (auf Personen bezogen)/ Yes (with regard to persons) Revision: Specify when the study has been performed, including beginning and end of the recruitment.	Kommentar: See above Kommentar: Is this an adequate translation of recruitment i.e. the process of inviting study participants?
15	Describe clinical [...] selection place).	Non	-	-	-	
16	Specify the number of participants who have met the inclusion criteria and who have undergone the index test and/ or the reference method. Give reasons why participants have not undergone one or the other test (a flow diagram is absolutely recommended).	B	O: strongly; FT: unbedingt, BT: absolutely Rückübersetzung ist an und für sich korrekt, aber dt. Version in diesem Falle inhaltlich zu stark formuliert The backward translation itself is correct. However, with in the forward translation (german version) the chosen language is	Änderung in „... (ein Flussdiagramm ist dringend zu empfehlen)“; plus FaÜ: Ist „dringend zu empfehlen“ mit „strongly recommended“ übersetzbar? Revision of German version into „... (ein Flussdiagramm ist dringend zu empfehlen)“;	Ja /Yes Revision: Specify the number of participants who have met the inclusion criteria and who have undergone the index test and/ or the reference method. Give reasons why participants have not undergone one or the	Kommentar: I Kommentar: Thank you for that strong language – it is not as strong in the 2003 STARD, („strongly recommended“) but this is something that we would like to change in a future revision of STARD

			stronger compared to the English original text passage.	plus QT: Can „dringend zu empfehlen“ be translated with „strongly recommended“?	other test (a flow diagram is strongly recommended).	
17	Specify the time period [...] performed since then.	Non	-	-	-	
18	Describe the distribution of the disease form (define the criteria) of the participants with the target condition (target criterion) and specify other diagnoses of participants without target condition (if applicable).	B	O: target condition, FT: Zielbedingung (Zielkriterium), BT: target condition (target criterion)	Revision of forward translation (German version): „Beschreiben Sie die Verteilung der Krankheitsausprägung (definieren Sie die Kriterien) bei den Teilnehmern mit der untersuchten Erkrankung; nennen Sie andere Diagnosen der Teilnehmer ohne diese Erkrankung“ Question to the original authors: When you write in the end of this issue "...; other diagnosis in participants without the target condition?", do you mean <i>all other diseases out of the participant's medical history</i> ?	Revision: Describe the distribution of the disease form (define the criteria) of the participants with the determined disease; name other diagnoses of participants without this disease.	Kommentar: Are these synonyms?
19	Present the results [...] the reference method.	Non	-	-	-	
20	Specify any unwanted results from the accomplishment of the index tests or the reference method.	A	O: adverse events, FT: unerwünschte Ereignisse, BT: unwanted results	FaU: Kann „unerwünschte Ereignisse“ auch mit „adverse events“ übersetzt werden? QT: Can „unerwünschte Ereignisse“ be translated with „adverse events“ in this context?	Ja /Yes Revision: Specify any adverse events from the accomplishment of the index tests or the reference method.	Kommentar: This should be „effects“, not „results“
21	Display [...] (e. g. 95% confidence interval).	Non	-	-	-	
22	Describe how [...] treated.	Non	-	-	-	
23	Display the estimation [...] if accomplished.	Non	-	-	-	
24	Display the estimation of test repeatability, if accomplished.		O: estimates, FT: Schätzungen, BT: estimation Gespräch der deutschen Autoren (M.S., J.-P.B., M.B., T.B.) über Frage, ob „estimates“ besser mit „Schätzwerte“ übersetzt werden müsste als - wie auch in den übrigen Punkten geschehen – mit „Schätzung“. Ergebnis der Diskussion: nein. Die deutschen Autoren geben den Originalautoren Recht mit Ihrem Kommentar. Aber sagen, dass der Sachverhalt im Deutschen keinen inhaltlich relevanten Unterschied macht („Schätzung“ oder „Schätzwert“.)	Discussion of the German authors (M.S., J.-P.B., M.B., T.B.) about the question whether “estimates” should better be translated with “Schätzwerte” and not with “Schätzung”. Conclusion: no. The German authors think the original authors are right with their comment. However, they also think that there is no relevant difference between “Schätzwert” and “Schätzung”.		Kommentar: Should be „estimates“ (results)
25	Discuss the clinical [...] survey results.	Non	-	-	-	

Abbildung 2: Harmonisierungsfassung der zertifizierten Rückübersetzung der deutschen STARD-Fassung ins Englische

7. Kognitives Debriefing

Nach der Harmonisierung sollte ein kognitives Debriefing erfolgen. Hierzu wird die Übersetzung einer Gruppe von 5-8 (muttersprachlichen) Personen vorgelegt, welche die Übersetzung auf Verständlichkeit überprüfen. Bei diesen Debriefern sollte es sich um Personen handeln, welche die Zielgruppe (Geschlecht, Alter, Bildungsgrad, etc.) des

übersetzten Instrumentes repräsentieren. Bei medizinischen Checklisten wird diese an 5-8 Patienten angewendet, um die Eignung der Übersetzung auf konzeptioneller Ebene zu testen. Durch das Verfahren können auch alternative Übersetzungsvorschläge durch die Debriefler gewonnen werden. Missverständnisse bei einzelnen Punkten der Übersetzung sollen aufgedeckt werden. Ihr Verbleib könnte zum Erheben ungenauer Daten führen.

Nach Genehmigung der deutschen Fassung der STARD-Checkliste durch die Originalautoren erfolgte das kognitive Debriefing. Hierzu wurde 13 wissenschaftlichen Mitarbeitern der Charité Berlin und der Deutschen Krankenhausgesellschaft die deutsche Übersetzung zur Beurteilung der inhaltlichen Verständlichkeit und des sprachlichen Ausdrucks vorgelegt.



CharitéCentrum für Anästhesiologie, OP-Management und Intensivmedizin

Charité - Universitätsmedizin Berlin | 13344 Berlin

Universitätsklinik für Anästhesiologie m. S. operative Intensivmedizin

Klinikdirektorin:
Univ.-Prof. Dr. C. Spies

Campus Virchow-Klinikum

Augustenburger Platz 1
13353 Berlin
Tel. +49 30 450 551001/02
Fax: +49 30 450 551909
anaesthesie-virchow-klinikum@charite.de

Campus Charité Mitte

Charitéplatz 1
10117 Berlin
Tel. +49 30 450 531012/52
Fax: +49 30 450 531911
anaesth@charite.de



<http://anaesthesieintensivmedizin.charite.de/>

Berlin, den 7. September 2010

STARD-Kriterien – Bitte um Unterstützung im ‚Cognitive Debriefing‘

Sehr geehrte Frau Kollegin, sehr geehrter Herr Kollege,

wissenschaftliche Forschungsergebnisse bedürfen einer standardisierten Darstellung, um Ergebnisse vergleichbar und nachvollziehbar zu machen. Hierzu wurden international in Abhängigkeit von Studientyp Standards in Checklistenform entwickelt, von denen die meisten bisher nur in Englischer Sprache vorliegen. Der Übersetzungsprozess selbst bedarf dabei ebenfalls eines einheitlichen und transparenten Verfahrens, um Fehlinterpretationen auszuschließen. Im Falle von Checklisten kann die unkritische Nutzung die Erhebung nicht valider Daten zur Folge haben; bei Behandlungsempfehlungen können sich womöglich fatale Folgen für Patienten ergeben.

Unsere Arbeitsgruppe hat eine deutsche Übersetzung des originalen ‚Standards for Reporting of Diagnostic Accuracy (STARD-) Statements‘ (Bossuyt PM et al. BMJ 2003) verfasst. Dabei handelt es sich um eine 25 Punkte umfassende Empfehlung zu Kriterien, die Berichterstattungen über Studien diagnostischer Genauigkeit mindestens enthalten sollten, um wissenschaftlichen Standards zu genügen (vergleichbar mit dem CONSORT-Statement zur Qualität der Berichterstattung über randomisierte kontrollierte Studien).

Diese STARD Kriterien lagen bisher nur in der englischen Originalfassung vor. Unsere Übersetzung erfolgte nach einem standardisierten Verfahren (ISPOR, Wild D et al. Value in Health 2005), welches zum Abschluss ein so genanntes ‚Cognitive Debriefing‘ beinhaltet. Bei solch einem Debriefing wird einer Gruppe potentieller Nutzer eines übersetzten Instruments die von den Originalautoren bereits genehmigte deutsche Fassung zur Begutachtung vorgelegt.

Wir wären Ihnen sehr dankbar, wenn Sie uns bei diesem Schritt behilflich sein könnten. Dazu finden Sie in der Anlage die deutsche Version der STARD-Kriterien. In der ersten Spalte hinter den jeweiligen Punkten möchten wir Sie bitten, (A) die inhaltliche Verständlichkeit und in der zweiten Spalte (B) die sprachlich Qualität mit den Schulnoten 1-6 zu bewerten. Am Ende des Dokuments ist optional Platz für zusätzliche frei formulierte Anmerkungen.

Ihre Unterstützung stellt den entscheidenden abschließenden Teil dieses Übersetzungsprozesses dar.

Vielen Dank dafür.

Dr. Philipp Breuer

Dr. Matthes Seeling

Prof. Dr. Claudia Spies

STARD-Kriterien - Checkliste zum Berichten von Studien diagnostischer Genauigkeit

Standardisierte Übersetzung des originalen 'Standards for Reporting of Diagnostic Accuracy (STARD)-statements' (Bossuyt PM et al. BMJ 2003)

Bitte bewerten Sie die einzelnen Punkte der deutschen Übersetzung der STARD-Kriterien mit den Schulnoten 1 bis 6

(1 = sehr gut, 2 = gut, 3 = befriedigend, 4 = ausreichend, 5 = mangelhaft, 6 = ungenügend);
und zwar (A) nach inhaltlicher Verständlichkeit und (B) nach sprachlichem Stil.

Die letzte Seite bietet die Möglichkeit für freie Kommentare.

Abschnitt und Thema	Punkt	Beschreibung	A Inhaltliche Verständlich keit	B Sprachlicher Stil
Titel, Zusammenfassung, Schlüsselwörter	1	Kennzeichnen Sie Ihren Artikel als eine Studie diagnostischer Genauigkeit (geben Sie als Schlüsselwörter die MeSH*-Begriffe „Sensitivität und Spezifität“ an).		
Einführung	2	Benennen Sie die wissenschaftlichen Fragestellungen oder Studienziele, wie zum Beispiel die Einschätzung diagnostischer Genauigkeit oder den Vergleich der Genauigkeit verschiedener Testverfahren oder zwischen unterschiedlichen Studiengruppen.		
Methoden:				
Teilnehmer	3	Beschreiben Sie die Studienpopulation: die Ein- und Ausschlusskriterien wie auch die Rahmenbedingungen und Orte der Datenerhebung.		
	4	Beschreiben Sie die Rekrutierung der Testteilnehmer: basierte diese auf vorhandenen Symptomen, auf Ergebnissen früherer Tests oder der Tatsache, dass Testteilnehmer den Index-Tests oder dem Referenzverfahren unterzogen worden waren? Anmerkung: Andere Rekrutierungsverfahren sind möglich.		
	5	Beschreiben Sie das Stichprobenauswahlverfahren für die Teilnehmer: wurden die Teilnehmer konsekutiv in die Studienpopulation eingeschlossen wie in den Auswahlkriterien unter 3 und 4 definiert? Wenn nicht, spezifizieren Sie, wie die Teilnehmer ausgewählt wurden.		
	6	Beschreiben Sie die Datenerhebung: wurde die Datenerhebung vor (prospektive Studie) oder nach (retrospektive Studie) der Durchführung des Index-Tests oder des Referenzverfahrens geplant?		
Testmethoden	7	Beschreiben Sie das Referenzverfahren und dessen Grundprinzip.		
	8	Beschreiben Sie die technischen Spezifika von Material und Methoden einschließlich		

		Art und Zeitpunkt der Messung, und/oder zitieren Sie Literaturquellen zu den Index-Tests und dem Referenzverfahren.		
	9	Nennen Sie die Definition und beschreiben Sie die Rationale der gewählten Einheiten, cut-offs und Kategorien der Messergebnisse von Index-Tests und Referenzverfahren.		
	10	Beschreiben Sie die Anzahl, die absolvierten Schulungsmaßnahmen und die Fachkompetenz der Personen, die den Index-Test und das Referenzverfahren durchgeführt und beurteilt haben.		
	11	Beschreiben Sie, ob die Beurteiler der Index-Tests und des Referenzverfahrens gegenüber den Ergebnissen des jeweils anderen Tests verblindet waren. Erläutern Sie jede weitere klinische Information, die den Beurteilern verfügbar war.		
Statistische Methoden	12	Beschreiben Sie die Methoden zur Berechnung oder zum Vergleich von Messergebnissen zur diagnostischen Genauigkeit und die verwendeten statistischen Verfahren, um die Messunsicherheit zu quantifizieren (zum Beispiel: 95% Konfidenzintervall).		
	13	Beschreiben Sie die Methoden zur Berechnung der Testreproduzierbarkeit, sofern durchgeführt.		
Ergebnisse:				
Teilnehmer	14	Geben Sie an, wann die Studie durchgeführt wurde inklusive Beginn und Ende der Rekrutierung.		
	15	Beschreiben Sie klinische und demographische Charakteristika der Studienpopulation (zum Beispiel Alter, Geschlecht, dargebotene Symptome, Komorbiditäten, aktuelle Therapien und Ort der Rekrutierung).		
	16	Nennen Sie die Anzahl der Teilnehmer, welche die Einschlusskriterien erfüllten und dem Index-Test und/oder dem Referenzverfahren unterzogen beziehungsweise nicht unterzogen wurden. Nennen Sie die Gründe, warum Teilnehmer dem einen oder anderen Test nicht unterzogen wurden (ein Flussdiagramm ist dringend zu empfehlen).		
Testergebnisse	17	Nennen Sie den Zeitraum zwischen Durchführung der Index-Tests und des Referenzverfahrens und berichten Sie über alle Behandlungen, die zwischenzeitlich erfolgten.		
	18	Beschreiben Sie die Verteilung der Krankheitsausprägung (definieren Sie die Kriterien) bei den Teilnehmern mit der untersuchten Zielerkrankung; nennen Sie andere Diagnosen der Teilnehmer ohne diese Erkrankung.		
	19	Stellen Sie die Ergebnisse der Index-Tests (einschließlich unbestimmbarer und fehlender Ergebnisse) gegen die Ergebnisse		

		des Referenzverfahrens in einer Kreuztabelle dar; bei kontinuierlichen Ergebnissen stellen Sie die Verteilung der Ergebnisse gegen die Ergebnisse des Referenzverfahrens dar.		
	20	Nennen Sie alle unerwünschten Ereignisse bei der Durchführung der Index-Tests oder des Referenzverfahrens.		
Schätzungen	21	Stellen Sie die Schätzung der diagnostischen Genauigkeit und die Messergebnisse zur statistischen Unsicherheit (zum Beispiel: 95% Konfidenzintervall) dar.		
	22	Beschreiben Sie, wie mit unbestimmbaren Ergebnissen, fehlenden Daten und Ausreißern des Index-Tests verfahren wurde.		
	23	Stellen Sie die Schätzung der Variabilität der diagnostischen Genauigkeit zwischen Beurteilern, Studienzentren oder Subgruppen von Teilnehmern dar, sofern durchgeführt.		
	24	Stellen Sie die Schätzung der Testreproduzierbarkeit dar, sofern erfolgt.		
Diskussion:	25	Diskutieren Sie die klinische Anwendbarkeit der Studienergebnisse.		

*MeSH Medical Subject Headings (United States - National Library of Medicine – National Institutes of Health)

Freie Kommentare		

Abbildung 4: Debriefing Checkliste

8. Review des kognitiven Debriefings und Finalisierung

In der Nachschau werden die Ergebnisse des kognitiven Debriefings mit der Originalversion verglichen unter Berücksichtigung kultureller Besonderheiten der Zielsprache. Dies dient der Ausräumung etwaiger im kognitiven Debriefingprozess gefundener Unklarheiten oder Missverständlichkeiten der Übersetzung und so der abschließenden Optimierung der Übersetzung. Dieses sollte unter der Leitung des Projektmanagers erfolgen. Wird nach Überprüfung von Anmerkungen aus dem Debriefingprozess und Berücksichtigung jeglicher als relevant erkannter Bedenken, zwischen Projektmanager, Übersetzern und Originalautoren Einvernehmen erzielt, so gilt das Projekt als finalisiert.

Die ‚key in-country‘- Repräsentanten führten das Review des kognitiven Debriefings mit dem Ziel der Finalisierung durch. Es wurden alle Einzelanmerkungen der Debriefe, und sämtliche Einzelbenotungen, die abgerundet die Note „befriedigend“ oder schlechter erhielten, überprüft.

9. Korrekturlesen

Hierzu existieren keine expliziten Angaben der TCA-Gruppe. Trotzdem bestand Einigkeit darüber diesen Punkt in die Leitlinien mit aufzunehmen. Das Korrekturlesen sollte durch den Projektmanager und/oder einen Lektor erfolgen, um kleinere eventuell noch nicht korrigierte Fehler aufzufinden.

Der Finalisierung folgte das Korrekturlesen der endgültigen Fassung. Dieses wurde von drei der vier ‚key in-country‘ - Repräsentanten vorgenommen.

10. Abschlussbericht

Dieser Bericht soll den Übersetzungsprozess erläutern und Beweggründe für die Wortwahl bei bestimmten Übersetzungen nennen, um auch späteren Übersetzungen oder Weiterentwicklungen bestehender Instrumente aus diesem Bereich als Grundlage zur Verfügung zu stehen, auch im Hinblick auf ein dann wieder erforderliches Harmonisierungsverfahren.

Insbesondere mit den Abschnitten ‚Methoden‘ und ‚Ergebnisse‘ entspricht die vorliegende Dissertation inhaltlich dem genannten Abschlussbericht.

3. Ergebnisse

Arbeitsaufwand des Übersetzungsprozesses

Der Arbeitsaufwand der beteiligten Personen bei den einzelnen Arbeitsschritten ist nach Anzahl der dazu notwendigen Treffen und in ihrem zeitlichen Aufwand in nachfolgender Tabelle dargestellt.

Tabelle 2 Arbeitsaufwand des Übersetzungsprozesses

Arbeitsschritte	Personen (n)	Konferenzen (n)	Arbeitszeit (Stunden)
Vorbereitung	3	1	12
Vorwärtsübersetzung	4	1	20
Konsensusfindung	4	4	24
Rückübersetzung	1* Übersetzerin	-	im Auftrag, Zeit nicht ermittelt
Review der Rückübersetzung	4	1	8
Harmonisierung	4	2	8
Kognitives Debriefing	13**	2	Zeit nicht ermittelt
Review des kognitiven Debriefings	4	2	16
Finalisierung	4	1	4
Korrekturlesen	3	-	5
Abschlussbericht	1	-	80
Gesamt	-	14	177

*Diplomierte Übersetzerin, **Debriefer (eigene Klinik und Deutsche Krankenhausgesellschaft)

Diskrepanzen zwischen Original und Rückübersetzung und Harmonisierung

Ausgang für den nach der ISPOR definierten Harmonisierungsprozess stellten die Ergebnisse aus der Beurteilung der Rückübersetzung im Vergleich mit dem Original dar. Die Originalautoren sahen nach Begutachtung der ins Englische zurückübersetzten Deutschen Version Korrekturbedarf bei den ‚STARD-Punkten‘ 1, 3, 4, 5, 14, 16, 18, 20 und 24. Die Punkte 1, 3, 4, 14, 20 und 24 wurden eher als technische Übersetzungsfehler gewertet, die Punkte 5, 16 und 18 wurden als inhaltliche Unterschiede ausgemacht (Tabelle 3).

Tabelle 3 Diskrepanzen zwischen Originaltext und Rückübersetzung

Kategorie	N	%
technische Beanstandungen	6	24
konzeptionelle Beanstandungen	3	12
keine Beanstandungen	16	64

Diskrepanzen technischer Art beinhalten sprachliche oder orthografische Unterschiede oder Änderungsvorschläge bezüglich des Sprachstils, Diskrepanzen konzeptioneller Natur beinhalten sprachlich inhaltlich relevante Diskrepanzen zwischen Original und Rückübersetzung

Zu den genannten Punkten war der Diskussions- und Entscheidungsprozess mit abschließender Lösung wie folgt:

Zu Punkt 1: „Identify your article as survey of diagnostic accuracy (recommend the MeSH* terms “sensitivity and specificity” as key words).”

Die Originalautoren bemängelten den Begriff „survey“ als Rückübersetzung des Begriffs Studie. Die Übersetzerin stimmte in diesem Zusammenhang zu, dass man den deutschen Begriff „Studie“ mit „study“ übersetzen kann. Dies gilt auch für den in anderen Punkten gewählten Begriff „survey“; er wurde in „study“ abgeändert und von den Originalautoren wie auch der Übersetzerin akzeptiert.

Zu Punkt 3: „Describe the survey population: inclusion and exclusion criteria as well as general conditions and places of data collection.”

Die Originalautoren gaben zu Bedenken, dass der Begriff „general conditions“ nicht deckungsgleich mit „setting“ sei. Der Vorschlag anstatt des Begriffs „Rahmenbedingung“ den Begriff „setting“ zu wählen und somit eine wortgetreue Übersetzung des Originals zu haben wurde von der Übersetzerin akzeptiert.

Zu Punkt 4: Die Originalautoren gaben zu bedenken, dass die englische Fassung „presenting symptoms“ und unsere Version „existing symptoms“ nicht deckungsgleich sei. Die Übersetzerin stimmte zu, dass die von uns vorgeschlagene Version „vorhandene Symptome“ auch mit „presenting symptoms“ übersetzt werden kann.

Zu Punkt 5: Bei diesen Punkt wurde erstmals eine inhaltliche Differenz der Übersetzung deutlich und von den Originalautoren angemahnt. Diese wählten den Begriff „consecutive series“, welcher durch die ‚key in-country‘- Repräsentanten mit „fortlaufende Aufnahme“ übersetzt wurde. Die Rückübersetzung lautete „ongoing acceptance“. Zur Korrektur schlugen die ‚key in-country‘- Repräsentanten hier „konsekutiver Einschluss“ vor.

Zu Punkt 14: Die Originalautoren waren sich nicht sicher, ob der Originalbegriff „recruitment“, durch die Rückübersetzung „selection“ richtig widerspiegelt wird.

Die Übersetzerin stimmte unserem Vorschlag zu, „recruitment“ mit „Rekrutierung von Teilnehmern“ zu übersetzen.

Zu Punkt 16: Der Originalbegriff „strongly recommended“ wurde mit „dringend empfohlen“ übersetzt. Die Rückübersetzung lautete „absolutely recommended“. Die englische Fassung war den Originalautoren zu stark, obwohl sie bei einer zukünftigen Revision der englischen Originalversion der STARD-Kriterien einen solch starken Begriff durchaus vorzögen.

Die Übersetzerin stimmte dem Lösungsvorschlag zu, dass man „strongly recommended“ mit „dringend empfohlen“ übersetzen könne.

Zu Punkt 18: Von den Originalautoren wurde der Begriff „target condition“ verwandt, dieser wurde von uns mit „Zielbedingung (Zielkriterium)“ übersetzt und mit „target condition (target criterion)“ zurückübersetzt. Die ‚key in-country‘- Repräsentanten beschlossen eine inhaltliche Revision der Vorwärtsübersetzung wie folgt: „Beschreiben Sie die Verteilung der Krankheitsausprägung (definieren Sie die Kriterien) bei den Teilnehmern mit der untersuchten Erkrankung; nennen Sie andere Diagnosen der Teilnehmer ohne diese Erkrankung“ und erbaten eine Erklärung für den Satz „...other diagnoses in participants without the target condition?“ ob sämtliche Krankheiten in der Anamnese des Teilnehmers gemeint seien.

Zu Punkt 20: Der Originaltext verwendet den Begriff „adverse events“, die Übersetzung lautete „unerwünschte Ereignisse“, der Rückübersetzungsvorschlag „unwanted results“.

Die Originalautoren gaben zu bedenken, dass es sich hierbei um „effects“ und nicht „results“ handele. Dem Vorschlag „unerwünschte Ereignisse“ mit „adverse events“ zu übersetzen stimmte die Übersetzerin zu.

Zu Punkt 24: Das Original verwendet hier den Begriff „estimates“, die Vorwärtsübersetzung lautete „Schätzungen“, die Rückübersetzung „estimation“. Die Originalautoren waren der Meinung, dass der korrekte Begriff „estimates“ im Sinne von „results“ und nicht „estimation“ lauten müsste.

Hierzu erfolgte eine Beratung der ‚key in-country‘- Repräsentanten über die Frage, ob „estimates“ besser mit „Schätzwerte“ übersetzt werden müsste als - wie auch in den übrigen Punkten geschehen - mit „Schätzung“. Im Ergebnis der Diskussion Gaben die ‚key in-country‘- Repräsentanten den Originalautoren Recht mit ihrem Kommentar. Schlussendlich war man jedoch der Auffassung, dass der Sachverhalt im Deutschen keinen inhaltlich relevanten Unterschied macht („Schätzung“ oder „Schätzwert“).

Die aus diesen Ergebnissen gezogenen Schlüsse und entsprechend als notwendig erachteten Revisionen wurden dem Originalautor erneut vorgelegt, ohne dass dieser Beanstandungen hatte.

Kognitives Debriefing

Tabelle 4: Benotung im kognitiven Debriefing

Item	Median	Min-Max		Item	Median	Min-Max
1a	3,00	1,00-4,00		14a	1,00	1,00-3,00
1b	3,00	1,00-5,00		14b	1,00	1,00-2,00
2a	2,00	1,00-3,00		15a	1,00	1,00-3,00
2b	2,00	1,00-4,00		15b	1,00	1,00-2,00
3a	1,50	1,00-3,00		16a	1,00	1,00-3,00
3b	1,50	1,00-3,00		16b	2,00	1,00-3,00
4a	2,00	1,00-4,00		17a	1,00	1,00-2,00
4b	2,00	1,00-4,00		17b	1,00	1,00-3,00
5a	2,00	1,00-3,00		18a	2,00	1,00-5,00
5b	2,00	1,00-4,00		18b	2,00	1,00-5,00
6a	1,00	1,00-6,00		19a	1,00	1,00-3,00
6b	1,00	1,00-2,00		19b	2,00	1,00-3,00
7a	1,00	1,00-3,00		20a	1,00	1,00-2,00
7b	1,00	1,00-2,00		20b	1,00	1,00-2,00
8a	2,00	1,00-3,00		21a	1,00	1,00-3,00
8b	2,00	1,00-3,00		21b	1,00	1,00-3,00
9a	2,00	1,00-3,00		22a	1,00	1,00-3,00
9b	2,00	1,00-3,00		22b	1,00	1,00-2,00
10a	1,00	1,00-3,00		23a	2,00	1,00-2,00
10b	1,00	1,00-3,00		23b	2,00	1,00-4,00
11a	1,00	1,00-3,00		24a	1,50	1,00-3,00
11b	2,00	1,00-3,00		24b	1,50	1,00-4,00
12a	1,00	1,00-2,00		25a	1,00	1,00-2,00
12b	2,00	1,00-3,00		25b	1,00	1,00-2,00
13a	2,00	1,00-2,00				
13b	2,00	1,00-4,00				

Insgesamt wurde die inhaltliche Verständlichkeit im Median mit 1,00 (1,00 – 6,00), die sprachliche Verständlichkeit mit 2,00 (1,00 – 5,00) bewertet.

Wie in den Methoden zuvor definiert wurden alle Items, die von den Debriefern mit 3 oder schlechter benotet worden waren einer strengen Prüfung durch die Arbeitsgruppe unterzogen und in einer entsprechende Revision konsentiert. Selbst bei den besser benoteten Items wurde in Falle einer zusätzlichen freien Kommentierung durch die Debriefere eine eventuelle Revision der betreffenden Items geprüft. Schlussendlich waren alle nötigen Revisionen allerdings ausschließlich sprachlicher Natur. Diese Änderungen sind aus der Endversion im Abgleich mit der Debriefing- Version (s. unter „Methoden“) zu ersehen.

Deutsche Endversion der STARD-Kriterien

Die Vorlage dieser Version inklusive der Ergebnisse des Debriefings bei dem Originalautor ergab keine Beanstandungen. Auch das Korrekturlesen (ISPOR- Arbeitsschritt 9) ergab keinen Korrekturbedarf.

Abschnitt und Thema	Punkt	Beschreibung
Titel, Zusammenfassung, Schlüsselwörter	1	Kennzeichnen Sie Ihren Artikel als eine Studie diagnostischer Genauigkeit (geben Sie als Schlüsselwörter die MeSH*-Begriffe „Sensitivität und Spezifität“ an).
Einführung	2	Benennen Sie die wissenschaftlichen Fragestellungen oder Studienziele, wie zum Beispiel die Einschätzung diagnostischer Genauigkeit oder den Vergleich der Genauigkeit verschiedener Testverfahren oder zwischen unterschiedlichen Studiengruppen.
Methoden:		
Teilnehmer	3	Beschreiben Sie die Studienpopulation: die Ein- und Ausschlusskriterien wie auch die Rahmenbedingungen und Orte der Datenerhebung.
	4	Beschreiben Sie die Rekrutierung der Testteilnehmer: basierte diese auf vorhandenen Symptomen, auf Ergebnissen früherer Tests oder der Tatsache, dass Testteilnehmer den Index-Tests oder dem Referenzverfahren unterzogen worden waren? Anmerkung: Andere Rekrutierungsverfahren sind möglich.
	5	Beschreiben Sie das Stichprobenauswahlverfahren für die Teilnehmer: wurden die Teilnehmer konsekutiv in die Studienpopulation eingeschlossen wie in den Auswahlkriterien unter 3 und 4 definiert? Wenn nicht, spezifizieren Sie, wie die Teilnehmer ausgewählt wurden.
	6	Beschreiben Sie die Datenerhebung: wurde die Datenerhebung vor (prospektive Studie) oder nach (retrospektive Studie) der Durchführung des Index-Tests oder des Referenzverfahrens geplant?
Testmethoden	7	Beschreiben Sie das Referenzverfahren und dessen Grundprinzip.
	8	Beschreiben Sie die technischen Spezifika von Material und Methoden einschließlich Art und Zeitpunkt der Messung, und/oder zitieren Sie Literaturquellen zu den Index-Tests und dem Referenzverfahren.
	9	Nennen Sie die Definition und beschreiben Sie die Rationale der gewählten Einheiten, cut-offs und Kategorien der Messergebnisse von Index-Tests und Referenzverfahren.
	10	Beschreiben Sie die Anzahl, die absolvierten Schulungsmaßnahmen und die Fachkompetenz der Personen, die den Index-Test und das Referenzverfahren durchgeführt und beurteilt haben.
	11	Beschreiben Sie, ob die Beurteiler der Index-Tests und des Referenzverfahrens gegenüber den Ergebnissen des jeweils anderen Tests verblindet waren. Erläutern Sie jede weitere klinische Information, die den Beurteilern verfügbar war.
Statistische Methoden	12	Beschreiben Sie die Methoden zur Berechnung oder zum Vergleich von Messergebnissen zur diagnostischen Genauigkeit und die verwendeten statistischen Verfahren, um die Messunsicherheit zu quantifizieren (zum Beispiel: 95% Konfidenzintervall).
	13	Beschreiben Sie die Methoden zur Berechnung der Testreproduzierbarkeit, sofern durchgeführt.
Ergebnisse:		
Teilnehmer	14	Geben Sie an, wann die Studie durchgeführt wurde inklusive Beginn und Ende der Rekrutierung.
	15	Beschreiben Sie klinische und demographische Charakteristika der Studienpopulation (zum Beispiel Alter, Geschlecht, dargebotene

		Symptome, Komorbiditäten, aktuelle Therapien und Ort der Rekrutierung).
	16	Nennen Sie die Anzahl der Teilnehmer, welche die Einschlusskriterien erfüllten und dem Index-Test und/oder dem Referenzverfahren unterzogen beziehungsweise nicht unterzogen wurden. Nennen Sie die Gründe, warum Teilnehmer dem einen oder anderen Test nicht unterzogen wurden (ein Flussdiagramm ist dringend zu empfehlen).
Testergebnisse	17	Nennen Sie den Zeitraum zwischen Durchführung der Index-Tests und des Referenzverfahrens und berichten Sie über alle Behandlungen, die zwischenzeitlich erfolgten.
	18	Beschreiben Sie die Verteilung der Krankheitsausprägung (definieren Sie die Kriterien) bei den Teilnehmern mit der untersuchten Zielerkrankung; nennen Sie andere Diagnosen der Teilnehmer ohne diese Erkrankung.
	19	Stellen Sie die Ergebnisse der Index-Tests (einschließlich unbestimmbarer und fehlender Ergebnisse) gegen die Ergebnisse des Referenzverfahrens in einer Kreuztabelle dar; bei kontinuierlichen Ergebnissen stellen Sie die Verteilung der Ergebnisse gegen die Ergebnisse des Referenzverfahrens dar.
	20	Nennen Sie alle unerwünschten Ereignisse bei der Durchführung der Index-Tests oder des Referenzverfahrens.
Schätzungen	21	Stellen Sie die Schätzung der diagnostischen Genauigkeit und die Messergebnisse zur statistischen Unsicherheit (zum Beispiel: 95% Konfidenzintervall) dar.
	22	Beschreiben Sie, wie mit unbestimmbaren Ergebnissen, fehlenden Daten und Ausreißern des Index-Tests verfahren wurde.
	23	Stellen Sie die Schätzung der Variabilität der diagnostischen Genauigkeit zwischen Beurteilern, Studienzentren oder Subgruppen von Teilnehmern dar, sofern durchgeführt.
	24	Stellen Sie die Schätzung der Testreproduzierbarkeit dar, sofern erfolgt.
Diskussion:	25	Diskutieren Sie die klinische Anwendbarkeit der Studienergebnisse.

*MeSH Medical Subject Headings (United States - National Library of Medicine – National Institutes of Health)

Abbildung 5: Deutsche Endfassung der STARD-Checkliste

4. Diskussion

Die wesentlichen Ergebnisse dieser Arbeit sind:

- i. Die ISPOR-Richtlinien ließen sich schlüssig auf den Übersetzungsprozess des STARD-Statements anwenden
- ii. Innerhalb des Übersetzungsprozesses kam es in 12% der Fälle zu einem Korrekturbedarf in konzeptioneller Hinsicht
- iii. Die vorliegende Deutsche Version des STARD-Statements wurde von den Originalautoren autorisiert

Nach bestem Wissen sind im deutschsprachigen Raum bisherige Übersetzungen von in Englisch verfassten Richtlinien zum präzisen Berichten klinischer Studien ohne einen definierten Übersetzungsprozess erfolgt [54-56]. Die vorliegende Übersetzung des STARD-Statements erfolgte nach den Standards der ‚Translation and Cultural Adaptation‘ (TCA)-group der ISPOR.

Mit der vorliegenden Übersetzung der STARD-Kriterien wird Personen, welche über die Durchführung einer Studie diagnostischer Genauigkeit berichten, als auch denen, die einen solchen Bericht lesen ein Instrument zur Verfügung gestellt, welches sie in die Lage versetzt die Qualität der erhobenen Daten zu beurteilen. In Form einer Checkliste stehen nunmehr 25 Punkte in der deutschen Sprache zur Verfügung, die nach Meinung der Originalautoren einen Mindeststandard für das Durchführen und Berichten über Diagnosestudien darstellen.

Gemäß den Guidelines der Translation and Cultural Adaptation group standen die Autoren mit den Originalautoren in Verbindung, um Unklarheiten bezüglich Formulierungen oder Inhalt zu beseitigen. Nach Auffassung der Autoren konnte so eine authentische deutsche Übersetzung gefertigt werden.

Techniken zur sprachlichen Adaptation medizinischer Schriftinstrumente

Die Qualität von Daten aus übersetzten Instrumenten hängt nicht zuletzt von der Qualität des Übersetzungsprozesses ab [8]. Im Jahre 1999 wurde die TCA-Gruppe der ISPOR gegründet. Zum damaligen Zeitpunkt gab es keine standardisierten Verfahren zum Übersetzen medizinischer Fachtexte oder Messinstrumente. Es bestanden viele Herangehensweisen, von

der ausschließlichen Vorwärtsübersetzung durch einen einzelnen Übersetzer zu Verfahren mit mehreren Rückübersetzungen. Die TCA-Gruppe verglich zur Bestandsaufnahme angewandter Verfahren 12 bestehende Leitlinien zur Übersetzung und kulturellen Anpassung der unter Punkt 2, Methodik aufgeführten Gesellschaften.

Bei dieser Durchsicht fanden sich Unterschiede hinsichtlich der Konsistenz in der Terminologie als auch der Methodik. In der Praxis werden verschiedene Methoden verwandt, um ähnliche Aufgaben zu erfüllen. Ein standardisiertes Vorgehen bei Übersetzungen ist jedoch unerlässlich, damit Ergebnisse vergleichbar und konsistent sind. Zudem werden unterschiedliche Terminologien für den gleichen Sachverhalt benutzt, was die Verständlichkeit erschwert. So werden die Begriffe „kognitives Debriefing“ und „Pilot-Testung“ teilweise synonym gebraucht [7]. Das kognitive Debriefing meint häufig die Anwendung eines neu übersetzten Messinstrumentes an einer Gruppe, welche nur 5 oder 6 Patienten umfasst [7]. Unter einer Pilot-Testung wird häufig die Anwendung eines neu übersetzten Messinstrumentes an einer Gruppe von 30-40 Patienten verstanden [7].

Des Weiteren machte die ISPOR ein Mangel an Methodenkonsistenz aus. Manche übersetzten Instrumente basierten auf Vorgehensweisen mit Vorwärtsübersetzungen und einer oder mehreren Rückübersetzungen [7], andere verzichteten auf Rückübersetzungen als Kontrollinstrument der Vorwärtsübersetzung [7].

Die ISPOR empfiehlt die Durchführung einer Rückübersetzung. Die vorliegende deutsche Version wurde durch eine Übersetzerin in die englische Sprache zurück übersetzt und durch die Originalautoren begutachtet. Drei der 25 Punkte wiesen dabei konzeptionelle Differenzen auf, welche ohne eine Rückübersetzung unberücksichtigt geblieben wären. Insgesamt ließ sich die STARD-Checkliste gut in die deutsche Sprache übertragen.

Larsson et al. [21] übertrugen die so genannte CAM-ICU (Confusion assessment method for the intensive care unit) aus dem englischen Original von Ely et al. [22] in die schwedische Sprache und hielten sich dabei an die Leitlinien der TCA-Gruppe und fertigten entsprechend auch eine Rückübersetzung. Der Übersetzungsprozess verlief nach Aussagen der Autoren ebenfalls relativ reibungsfrei, allerdings bereitete das Wort „Delirium“ und das Wort „Confusion“ Schwierigkeiten, da im schwedischen Sprachgebrauch diese mit Alkoholmissbrauch assoziiert sind und als beleidigend empfunden werden [21]. Hier war ein zusätzlicher Zeitaufwand erforderlich Angehörigen die Bedeutung dieser Begriffe im Zusammenhang der Intensivmedizin zu erläutern bevor diese Begriffe verstanden und akzeptiert wurden.

Selbst in Ländern mit ähnlichem geschichtlichem und kulturellem Hintergrund können unterschiedliche Gepflogenheiten die Übertragung eines Instruments in die andere Sprache schwierig gestalten. So übertrugen Boedeker et al. [23] die „Neighborhood Environment Wakability Scale“ (NEWS) [24] aus der englischsprachigen amerikanischen Originalfassung ins Deutsche. Mit diesem Fragebogen soll das subjektive Empfinden der „Begehbarkeit“ der eigenen Wohnumgebung, d.h. Bewegungsfreundlichkeit, speziell beim Zuzußgehen ermittelt werden.

Dabei ergaben sich unter anderem Probleme bei der Übertragung von Größenvorstellungen wie „football field“ (= amerikanisches Fußballfeld mit einer Feldgröße von 120 – 140 yards) auf Fußballfeld (europäischer Fußball mit einer Feldgröße von 100-130 yards), da man annahm, dass der Nutzer in Deutschland eher eine Vorstellung von der Größe eines Fußballplatzes hat, als von der Größe eines Platzes auf dem American football gespielt wird.

Die Frage nach der Gehfähigkeit von 100 yards wurde in 100 Meter übertragen, obwohl dies etwa 109 yards entspricht, aber eine Person in Deutschland sicher mit geraden Meterangaben mehr anzufangen weiß, als mit Angaben in yards.

Auch das Finden einer Übersetzung von „convenience store“ warf Probleme auf, da dieser in Deutschland kein typisches Geschäftsmodell darstellt.

Noch schwieriger gestaltet sich zuweilen das Übersetzen in Sprachen von Ländern mit sehr verschiedenem geschichtlichem und kulturellem Hintergrund.

Die Wichtigkeit einer Rückübersetzung zeigte sich eindrücklich bei der Übertragung des MOS-SSS („Medical Outcomes Study Social Support Survey“) [25] vom Englischen ins Chinesische [26]. Die Übersetzung folgte den Empfehlungen zu Übersetzungen von Brislin [27]. Hier wurde zunächst eine Übersetzung des englischen Originals ins Chinesische durch eine bilinguale Krankenschwester gefertigt. Diese Fassung wurde durch einen Laien mit der Muttersprache Chinesisch ohne Englischkenntnisse auf Verständlichkeit und Grammatik überprüft.

Die Chinesische Version wurde nun durch eine bilinguale Krankenschwester ohne Kenntnis der Englischen Originalfassung in das Englische zurückübersetzt. Schließlich erfolgte der Vergleich von Englischem Original und Englischer Übersetzung der chinesischen Version.

Hier zeigten sich grammatische, semantische und kulturelle Unterschiede. Bestimmte grammatische Wendungen existierten im Chinesischen nicht, und die Rückübersetzung ins Englische erhielt eine andere Bedeutung. Vor diesem Hintergrund ist eine Übersetzung mit

Rückübersetzung sicherlich sinnvoll um Bedeutungsunterschiede aufzudecken. Schwierig ist auch die Übertragung unterschiedlicher Gewohnheiten in verschiedenen Kulturkreisen. Bestimmte im europäisch- nordamerikanischen Kulturkreis übliche Handlungen und dazugehörige Beschreibungen sind im chinesischen Kulturkreis nicht gebräuchlich.

So kann die Übertragung der Frage „haben Sie Schwierigkeiten beim Essen mit Messer und Gabel?“ zwar vermutlich einfach gelingen, die Beantwortung jedoch das Ergebnis verfälschen, wenn im betreffenden Kulturkreis das Essen mit Messer und Gabel nicht üblich ist und somit für zumindest einen Teil der dort lebenden Menschen ungewohnt ist und von diesen mit „ja“ beantwortet würde [10].

Zu berücksichtigen ist auch die unterschiedliche Bedeutung von Begriffen in unterschiedlichen Kulturkreisen. So hat der Begriff der „Familie“ in Europa sicherlich eine andere Bedeutung als im islamischen oder fernöstlichen Kulturkreis, und eine Fragestellung in einem Messinstrument wie „können Sie ihre Familie so oft sehen, wie sie wollen?“ kann zu schwer zu vergleichenden Antworten führen [10].

Sowohl McKenna als auch Lenderking zweifeln die Überlegenheit einer Rückübersetzung gegenüber anderen Verfahren an [28;29]. Sie favorisieren den „dual-panel-approach“ [30]. Hierbei sollte eine Übersetzung des Originaldokuments von idealerweise 5-7 Übersetzern gefertigt werden, welche sich aus unterschiedlichen Tätigkeitsbereichen rekrutieren sollten, um ein breites Spektrum an Sprachgewohnheiten zu repräsentieren.

Die hieraus konsentierete Übersetzung sollte dann von 5-7 Laien ohne Kenntnis des Originals auf Stil und Verständlichkeit beurteilt werden. Die Laien sollten in etwa dem Bildungsgrad der Benutzer des übersetzten Instrumentes entsprechen; um die Allgemeinverständlichkeit des übersetzten Instruments zu prüfen, wird auch die Beteiligung von Laien mit unterdurchschnittlichem Bildungsgrad empfohlen. Schließlich sollte diese Fassung an etwa 15-20 Personen, die der geplanten Zielgruppe angehören ausprobiert werden. In ihrem Artikel geben McKenna et al. [28] zu bedenken, dass die TCA Gruppe den so genannten „dual-panel-approach“ bei ihrer Literaturdurchsicht und ihren Empfehlungen gänzlich unberücksichtigt ließ.

Nach McKenna [28] sollte eine Übersetzung von Personen gefertigt werden, die auch potentielle Nutzer des übersetzten Instrumentes sind. So sei der Gebrauch der themenüblichen Sprache gesichert [30]. Eine Rückübersetzung sei nicht erforderlich. Vielmehr sollte der „dual panel-approach“ genutzt werden [28].

Für Anpassungen von Messinstrumenten in anderen Ländern des gleichen Sprachraumes (Übertragung von mexikanischem Spanisch in argentinisches Spanisch) ist eine Übersetzung durch Laien nach McKenna [28] völlig ausreichend.

Außerdem würde eine Rückübersetzung die Kompetenz des oder der Übersetzer kompromittieren.

Gelegenheit zur Überprüfung, ob eine Rückübersetzung die Qualität einer Übertragung eines medizinischen Textes in eine andere Sprache erhöht oder ob die Dual-Panel- Methode inhaltliche und sprachliche Äquivalenz bietet, ergab sich, als neben der offiziellen Schwedischen Version des „Rheumatoid Arthritis Quality of Life“ [31] (RAQoL) Fragebogens [32] eine nicht autorisierte Schwedische Version des RAQoL [33] erschien, welche auf einer Rückübersetzung basierte, während die offizielle Übersetzung nach der Dual-Panel- Methode erfolgt war. Beim Vergleich beider Fassungen fanden sich bei dem 30 Punkte umfassenden Fragebogen 7 identisch und 23 unterschiedlich übersetzte Punkte. Die 23 unterschiedlich übersetzten Punkte wurden verschiedenen Gruppen im Hinblick auf Verständlichkeit und Wortwahl zur Beurteilung vorgelegt [34]. Eine Gruppe bestand aus zweisprachig aufgewachsenen Personen (n=23), eine andere Gruppe repräsentierte Laien (n=11), die dritte Gruppe bestand aus Patienten mit rheumatoider Arthritis (n=50). Hierbei entschieden sich die Laien und die Patienten bei den einzelnen übersetzten Punkten häufiger für die nach der Dual-Panel- Methode erstellte Fassung, während die zweisprachigen Personen häufiger der Version, die auf der Rückübersetzung basiert den Vorzug gab [34].

Unsere deutsche Fassung der STARD-Kriterien basiert auf einer Rückübersetzung und zielt auf Fachpublikum ab. Die Fassung wurde nicht durch Laien überprüft, sie wurde aber auch nicht für diese oder für Patienten erstellt.

Auch Lenderking [29] zweifelt den Aufwand der exakten Durchführung aller Schritte, der in den ISPOR-Kriterien zur richtlinienkonformen Übersetzung im Hinblick auf den dadurch zu erreichenden Qualitätsgewinn an. Auch er sieht nicht, dass eine Rückübersetzung zwingend besser sein muss als eine sorgfältig gefertigte Übersetzung. Im Falle des Kognitiven Debriefings sollten empirische Studien die Reliabilität und Validität überprüfen, ob sich ein Unterschied ergibt, wenn dieser Punkt ausgelassen wird [29].

Kognitives Debriefing

Die Entwicklung von Fragebögen wirft immer die Frage auf, ob die Beantwortung der Fragen gute Daten liefert, das heißt, ob sie das messen, was sie messen sollen und damit reliable und valide Antworten liefern [35;36].

Bis Mitte der 1980er Jahre wurden als Evaluationsverfahren zumeist so genannte Pretests eingesetzt; ein Verfahren, bei dem der Fragebogen an einer Ministichprobe unter Feldbedingungen getestet wird. Üblich waren unterschiedlichste Pretestvarianten, was die Vielzahl unterschiedlichster Begriffe erklärt, mit der Autoren, die von ihnen beschriebene und empfohlene Vorgehensweise bezeichneten [37].

In der sozialwissenschaftlichen Literatur finden sich unterschiedliche Empfehlungen zur Durchführung von Pretests. Die empfohlene Fallzahl bei einer Stichprobe, das heißt die Anzahl der zu befragenden Personen variiert je nach Autor zwischen $N = 10$ und $N = 200$ [38-47].

Im angloamerikanischen Raum wurden in der Kognitionspsychologie [48] kognitive Techniken entwickelt mit Hilfe derer gezielt das Frageverständnis der Befragten beurteilt werden sollte. Im Unterschied zu den Pretestverfahren, wo Unklarheiten nur offenbar werden, wenn der Benutzer sie äußert, werden bei kognitiven Techniken, Benutzer aufgefordert, ihre Gedankengänge zu den einzelnen Punkten laut darzulegen, so dass Verständnisprobleme leichter aufgedeckt werden können [37].

Die von der ISPOR untersuchten Leitlinien zur Übersetzung und kulturellen Anpassung medizinischer Veröffentlichungen empfahlen ein kognitives Debriefing des neu übersetzten Instruments, es bestanden jedoch unterschiedliche Empfehlungen wie viele Personen an diesem Debriefing teilnehmen sollten.

In den Richtlinien der ISPOR wird ein Debriefingverfahren mit 5-8 Personen empfohlen. Die ISPOR schreibt jedoch nicht ausdrücklich vor, dass ein Debriefing mit mehr als 8 Personen unzulässig ist.

Radtke et al. [49] führten ein Debriefing mit 8 ärztlichen Mitarbeitern und 10 Pflegekräften durch, um die Verständlichkeit und Durchführbarkeit der von ihnen übersetzten „Intensive Care Delirium Screening Checklist“ (ICDSC) [50] zu evaluieren. Die ICDSC wurde als Routine Screening Instrument zur Erkennung des postoperativen Delirs entwickelt.

Die Verständlichkeit und die Durchführbarkeit wurden durchweg gut oder sehr gut beurteilt, nur bei jeweils einem der acht Items ergab sich einmal ein signifikanter Unterschied in der Beurteilung durch Ärzte bzw. einmal durch das Pflegepersonal.

Tavares MGS [51] führten sogar 2 Debriefingverfahren bei der Übertragung des „Asthma Control Scoring System“ (ACSS) [52] in Brasilianisches Portugiesisch durch.

Bei der ACSS handelt es sich um eine Checkliste zur Evaluation der Effektivität der Asthmakontrolle. Nach ihrer Übersetzung wurde diese einem potentiellen Benutzerkreis bestehend aus 5 Physiotherapeuten und 5 Pulmologen zur Überprüfung von Verständlichkeit und Eindeutigkeit der Aussagen vorgelegt. Punkte, welche nicht einen Score von mindestens 0.8 (0.8- 1.0 entspricht eindeutige Verständlichkeit der Aussage) erreichten, wurden überarbeitet und die angepasste Checkliste wurde in einem zweiten Debriefingverfahren wiederum jeweils 5 Physiotherapeuten und Pulmologen zur Beurteilung vorgelegt. Bei dem 2. Verfahren erreichten alle Punkte einen Score ≥ 0.8 .

Die Anzahl der bei unserem Debriefingverfahren involvierten Kollegen betrug 13. Durch die Einbeziehung von 13 Personen wird die Qualität der Übersetzung durch eine größere Anzahl von Verbesserungsvorschlägen leicht erhöht, ebenso wie die Wahrscheinlichkeit des Aufdeckens verbliebener Fehler im übersetzten Instrument.

Die Debriefler entstammten der Charité und der Deutschen Krankenhausgesellschaft. Somit handelt es sich um die so genannte Zielpopulation, für welche die STARD-Kriterien erstellt und erstmals offiziell übersetzt wurden.

Bei dem durch uns durchgeführten kognitiven Debriefing fiel folgendes auf: Eine Reihe von Debriefern setzte zwar strenge Maßstäbe bei ihrer Beurteilung an, leider machten sie jedoch keine Verbesserungsvorschläge. Die schlechteste Einzelbenotung erhielt die inhaltliche Verständlichkeit des Punktes 6 mit der Note 6. Ein Verbesserungsvorschlag durch diesen Debriefler unterblieb leider.

Für den sprachlichen Stil erhielten Punkt 1 und Punkt 18 von je einem Debriefler die Note 5; leider wurde auch hier kein Verbesserungsvorschlag gemacht.

Es ist anzunehmen, dass das Benoten einfach schneller durchzuführen ist und eine schlecht benotete Übersetzung einen Hinweis darstellt, diesen Punkt der Übersetzung noch einmal zu überprüfen, ein Verbesserungsvorschlag vom Debriefler in Zeiten hoher Arbeitsdichte jedoch nicht erwartet werden kann.

Ein anderer Debriefler vergab für den sprachlichen Stil von Punkt 18 die Note 2, bewertet die inhaltliche Verständlichkeit mit Note 5.

Es fällt des Weiteren auf, dass besonders die ersten 6 Fragen relativ streng benotet wurden, und die Debriefler danach mehrheitlich Milde walten ließen. Nach der Note 6 in Punkt 6, vergab der Debriefler 30 mal Note 1.

Diese Tatsache dürfte einer gewissen „Fragenmüdigkeit“ geschuldet sein. Allerdings war durch uns beabsichtigt allen Debriefern den gleichen Fragebogen (mit gleicher Reihenfolge der Fragen) zur Verfügung zu stellen.

Ein Debriefler machte einige Verbesserungsvorschläge, für welche er sich selbst dann die Note 2 gab.

Die Qualität eines Debriefingverfahren ist nicht zuletzt von der Lust und dem Interesse der Debriefler sich mit dem debrieften Objekt zu befassen abhängig. Die Beurteilung des Sprachstils ist ebenfalls sehr subjektiv, ähnlich einer numerischen Analogskala, bei der Patienten die Stärke ihrer Schmerzen auf einer Skala von 0 (entsprechend: Kein Schmerz) bis 10 (entsprechend: Stärkster vorstellbarer Schmerz) angeben sollen [53]. Die subjektive Beurteilung des Schmerzes erschwert hier die Vergleichbarkeit der erhobenen Daten, wenn ein Patient seine durchschnittlichen Schmerzen mit Stufe 9-10 angibt, aber die Einnahme von Analgetika noch nicht für erforderlich hält, während ein anderer Patient in der Betrachtung des Arztes offensichtlich und trotz intensiver Schmerztherapie leidet, aber auf der Schmerzskala Stufe 1-2 angibt.

Trotzdem ist die Anwendung eines Debriefingverfahrens sinnvoll, da hiermit einer Gruppe von potentiellen Nutzern ein neues Instrument vorgestellt wird, und diese durch den Übersetzungsprozess unbeeinflusst, sprachliche Mängel eventuell aufdecken kann, die den ständig involvierten Personen sozusagen „betriebsblind“ entgangen sind.

Limitationen

Die vorliegende Arbeit weißt folgende Limitationen auf: Entgegen der Soll-Empfehlung der TCA-Gruppe der ISPOR ist die Rückübersetzerin nicht Muttersprachlerin der Originalsprache, sondern deutsche Muttersprachlerin. Diese Tatsache hat vermutlich jedoch nur einen geringen Einfluss auf die Qualität und die Übereinstimmung der übersetzten Fassung, da mehrere von ärztlichen Kollegen unabhängig gefertigte Vorwärtsübersetzungen konsentiert wurden. Von den Originalautoren wurden dann 9 von 25 Punkten beanstandet,

wobei es sich nur bei 3 von den 9 beanstandeten Punkten um inhaltliche Diskrepanzen handelte, die anderen 6 Punkte waren eher sprachlicher Natur.

Eine große Anzahl von Übersetzungen wurde nicht nach Rücksprache mit den Originalautoren oder ohne die Berücksichtigung von Übersetzungsleitlinien gefertigt [8]. Das Unterlassen kann hier gegebenenfalls zu Urheberrechtsverletzungen führen und in der Folge zu juristischen Konsequenzen. Die über die vorliegende Arbeit entstandene deutsche Version der STARD-Kriterien ist eine von den Originalautoren autorisierte und mit ihnen abgestimmte deutsche Übersetzung der originalen Version in Englischer Sprache.

Die Übersetzungsleitlinien der TCA-Gruppe wurden in ihrer Fassung aus dem Jahre 2005 zunächst für klinische Messinstrumente entwickelt [8]. Andere Gebiete, wie die Übersetzung von Symptomchecklisten oder die Anpassung von Instrumenten des gleichen Sprachraums an die kulturellen Gegebenheiten anderer Länder (Übertragung von in Spanien gesprochenem Spanisch in Spanisch, welches zum Beispiel in Argentinien üblich ist), sollten in späteren Treffen erörtert werden. Die Übersetzung der STARD-Checkliste ins Deutsche ist somit nicht primäres Ziel der TCA-Leitlinienempfehlungen. Die Qualität einer Übersetzung, und die Erhebung von Daten mit Messinstrumenten, welche aus einem unsystematischen Übersetzungsprozess stammen, sind jedoch zumindest fragwürdig.

5. Schlussfolgerung

Die vorliegende Arbeit zeigt, dass für die optimale sprachliche und inhaltliche Adaptation eines bereits bestehenden medizinischen Schriftinstrumentes wie das STARD-Statement an einen neuen Sprachraum der strukturierte Übersetzungsprozess der TCA schlüssig anwendbar ist. Die ISPOR-Richtlinien erscheinen somit für eine strukturierte Adaption von definierten Kriterien zur Berichterstattung von Studienergebnissen in andere Sprachen geeignet. In wie fern sich diese Methodik von anderen Übersetzungsprozessen unterscheidet ist Gegenstand künftiger Forschung.

Die vorliegende deutsche Version der STARD-Kriterien zum Berichten über Diagnosestudien stellt ein einfach zu handhabendes Instrument dar, um die Qualität eines Berichts über Studien dieser Art zu verifizieren und eigene Studien auf Vollständigkeit hinsichtlich aller Punkte, welche nach den STARD-Kriterien einen Mindeststandard im Hinblick auf qualitativ hochwertiges Berichten darstellen zu überprüfen.

Die deutsche Fassung soll die Verbreitung der STARD-Kriterien im deutschen Sprach- und Kulturraum fördern und ein Bewusstsein für die Notwendigkeit des Einhaltens bestimmter Mindestanforderungen an Studienbeschreibungen wecken.

6. Zusammenfassung

Hintergrund. Wissenschaftliche Forschungsergebnisse bedürfen einer standardisierten Darstellung, um nachvollziehbar und vergleichbar zu sein. Hierzu wurden in Abhängigkeit vom Studientyp offizielle Empfehlungen entwickelt, die im Original meist in Englisch vorliegen. So auch das so genannte ‘STAndards for Reporting of Diagnostic accuracy (STARD)-Statement’, eine 25 Kriterien umfassende Checkliste zur Berichterstattung von Diagnosestudien. Ein Angebot in anderen Sprachen, könnte die Berücksichtigung solcher Qualitätskriterien für wissenschaftliche Publikationen international steigern. Allerdings birgt jede Übersetzung die Gefahr von inhaltlichen Übertragungsfehlern in sich. Die ‚International Society for Pharmacoeconomics and Outcome Research‘ (ISPOR) hat zum Zwecke einer sicheren Adaptation in andere Sprachen Richtlinien vorgeschlagen.

Ziel der Arbeit war, über die Anwendung der ISPOR-Kriterien zur standardisierten Übersetzung eine durch die englischsprachigen Originalautoren autorisierbare Deutsche Version der STARD-Empfehlungen zu erstellen.

Methoden. Die Übersetzung des STARD-Statements erfolgte in Zusammenarbeit mit den Originalautoren nach den ISPOR-Arbeitsschritten 1. Vorbereitung, 2. voneinander unabhängige Vorwärtsübersetzungen, 3. Abstimmung, 4. professionelle Rückübersetzung, 5. Begutachtung der Rückübersetzung durch die Originalautoren und 6. Harmonisierung nach deren Kommentaren. 7. Kognitives Debriefing mit Bewertung der Empfehlungskriterien nach inhaltlicher Verständlichkeit und sprachlicher Qualität; nach entsprechender Revision (8.) erfolgte 9. die Finalisierung.

Ergebnisse. Die ISPOR-Richtlinien ließen sich schlüssig auf den Übersetzungsprozess des STARD-Statements anwenden. Der Aufwand umfasste personell 4 wissenschaftliche Mitarbeiter und eine diplomierte Übersetzerin, die kumulativ 177 Stunden an dem Projekt beschäftigt waren. Die Debriefing bewerteten die Empfehlungskriterien im Median (Min-Max) mit den Schulnoten 1,00 (1,00-6,00) (inhaltliche Verständlichkeit) und 2,00 (1,00-5,00) (sprachlicher Stil). Die Deutsche STARD-Version wurde von den Originalautoren autorisiert.

Schlussfolgerung. Die ISPOR-Richtlinien erscheinen für eine strukturierte Adaption von definierten Kriterien zur Berichterstattung von Studienergebnissen wie dem STARD-Statement in andere Sprachen geeignet.

7. Literatur

- 1 Begg C, Cho M, Eastwood S, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA* 1996;276:637-639.
- 2 Vandembroucke JP, von Elm E, Altman DG, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *Epidemiology* 2007;18:805-835.
- 3 von Elm E., Altman DG, Egger M, Pocock SJ, Gotsche PC, Vandembroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Prev Med* 2007;45:247-251.
- 4 Moher D, Cook DJ, Jadad et al. Assessing the quality of reports of randomised trials: implications for the conduct of meta-analyses. *Health Technol Assess* 1999;3:i-98.
- 5 Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *Quality of Reporting of Meta-analyses. Lancet* 1999;354:1896-1900.
- 6 Begg C, Cho M, Eastwood S, et al. [Improving the quality of reports on randomized controlled trials. Recommendations of the CONSORT Study Group]. *Rev Esp Salud Publica* 1998;72:5-11.
- 7 Wild D, Eremenco S, Mear I, et al. Multinational trials-recommendations on the translations required, approaches to using the same language in different countries, and the approaches to support pooling the data: the ISPOR Patient-Reported Outcomes Translation and Linguistic Validation Good Research Practices Task Force report. *Value Health* 2009;12:430-440.
- 8 Wild D, Grove A, Martin M, et al. 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* 2005;8:94-104.

- 9 Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ* 2003;326:41-44.
- 10 Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine* 2000;25:3186-3191.
- 11 Association of Test Publishers. Increasing the validity of adapted tests: myths to be avoided and guidelines for improving test adaptation practices. <http://www.testpublishers.org>.
- 12 Cull A, Sprangers M, Bjordal K, Aaronson N, West K, Bottomley A. EORTC Quality of life Study Group Translation Procedure; Brussels, EORTC, 1998.
- 13 Evidence Clinical and Pharmaceutical Resear: <http://www.evidence-cpr.com/>.
- 14 Bonomi AE, Cella DF, Hahn EA, et al. Multilingual translation of the Functional Assessment of Cancer Therapy (FACT) quality of life measurement system. *Qual Life Res* 1996;5:309-320.
- 15 Health Outcomes Group: Health Outcomes Group.
- 16 Szecket N, Medin G, Furlong WJ, Feeny DH, Barr RD, Depauw S. Preliminary translation and cultural adaptation of Health Utilities Index questionnaires for application in Argentina. *Int J Cancer Suppl* 1999;12:119-124.
- 17 Bullinger M, Alonso J, Apolone G et al. Translating health status questionnaires and evaluating their quality: the IQOLA approach. *J Clin Epidemiol* 1998;51:913-23.
- 18 Merkus MP, Dekker FW. *Kidney Disease Quality of Life-Short-Form: Translation Docum*; Amsterdam, The Netherlands, 1997.
- 19 Assessing health status and quality of life instruments: attributes and review: *Qual Life Res* 2002;11:193-205.
- 20 Sartorius N, Kuyken.W. Translation of health status instruments; in Orley J KW, (ed): *Quality of Life Assessment: International Perspectives*. Berlin Heidelberg, :Springer-Verlag, 1994.
- 21 Larsson C, Axell AG, Ersson A. Confusion assessment method for the intensive care unit (CAM-ICU): translation, retranslation and validation into Swedish intensive care settings. *Acta Anaesthesiol Scand* 2007;51:888-892.

- 22 Ely EW, Margolin R, Francis J, et al. Evaluation of delirium in critically ill patients: validation of the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU). *Crit Care Med* 2001;29:1370-1379.
- 23 Boedeker M, Bucksch J, Fuhrmann H. German Version of the Neighborhood Environment Walkability Scale (NEWS-G[ermany]) ; Bielefeld University, School of Public Health, Department of Prevention and Health Promotion, Germany, German Sport University Cologne, Institute of Sport Sociology , 2010.
- 24 Sallis JF. Neighborhood Environment Walkability Scale (NEWS). 2002.
- 25 Sherbourne CD, Stewart AL. The MOS social support survey. *Soc Sci Med* 1991;32:705-714.
- 26 Yu DS, Lee DT, Woo J. Psychometric testing of the Chinese version of the medical outcomes study social support survey (MOS-SSS-C). *Res Nurs Health* 2004;27:135-143.
- 27 Brislin JM. The most rewarding practice. *Am Pharm* 1985;NS25:7-8.
- 28 McKenna SP, Doward LC. The translation and cultural adaptation of patient-reported outcome measures. *Value Health* 2005;8:89-91.
- 29 Lenderking WR. Comments on the ISPOR Task Force Report on Translation and Adaptation of Outcomes Measures: guidelines and the need for more research. *Value Health* 2005;8:92-93.
- 30 Swaine-Verdier A, Doward LC, Hagell P, Thorsen H, McKenna SP. Adapting quality of life instruments. *Value Health* 2004;7 Suppl 1:S27-S30.
- 31 de Jong Z, van der Heijde D, McKenna SP, Whalley D. The reliability and construct validity of the RAQoL: a rheumatoid arthritis-specific quality of life instrument. *Br J Rheumatol* 1997;36:878-883.
- 32 McKenna SP, Hedin PJ. Adapting the Rheumatoid Arthritis Quality of Life Instrument (RAQoL) for use in Sweden. *Scand J Rheumatol* 2003;32:320-322.
- 33 Eberhardt K, Duckberg S, Larsson BM, Johnson PM, Nived K. Measuring health related quality of life in patients with rheumatoid arthritis--reliability, validity, and responsiveness of a Swedish version of RAQoL. *Scand J Rheumatol* 2002;31:6-12.

- 34 Hagell P, Hedin PJ, Meads DM, Nyberg L, McKenna SP. Effects of method of translation of patient-reported health outcome questionnaires: a randomized study of the translation of the Rheumatoid Arthritis Quality of Life (RAQoL) Instrument for Sweden. *Value Health* 2010;13:424-430.
- 35 Fox C. Questionnaire development. *J Health Soc Policy* 1996;8:39-48.
- 36 McDowell I, Praught E. On the measurement of happiness. An examination of the Bradburn scale in the Canada Health Survey. *Am J Epidemiol* 1982;116:949-958.
- 37 Prüfer P, Rexroth, M. Verfahren zur Evaluation von Survey - Fragen: Ein Überblick;ZUMA-Arbeitsbericht. Mannheim, ZUMA Zentrum für Umfragen, Methoden und Analysen 1996, p 4.
- 38 Backstrom CH, Hursh G. New York, Wiley, 1963.
- 39 Boyd HW, Westfall R. *Marketing Research. Text and Cases*; Homewood, Ill, R. D. Irwin, 1964.
- 40 Schrader A. *Einführung in die empirische Sozialforschung. Ein Leitfaden für die Planung, Durchführung und Bewertung von nicht-experimentellen Forschungsprojekten*; Stuttgart, Kohlhammer, 1971.
- 41 Elliott K, Christopher M. *Research Methods in Marketing*; London, Holt,Rinehart & Winston, 1973.
- 42 Karmasin F, Karmasin W. *Einführung in die Methode und Probleme der Umfrageforschung*. Wien, Hermann Böhlau Nachf., 1977.
- 43 Wellenreuther M. *Grundkurs: Empirische Forschungsmethoden Für Pädagogen, Psychologen, Soziologen* . Königstein, Athenäum, 1982.
- 44 Sheatsley PB. *Questionnaire Construction and Item Writing*. New York, New York:: Academic Press, 1983.
- 45 Fowler FJ. *Survey research methods*. Beverly Hills, CA: Sage Publications, Inc., 1984.
- 46 Converse JM, Presser S. *Survey Questions. Handcrafting the Standardized Questionnaire*. Beverly Hills, CA, Sage Publications, Inc., 1986.
- 47 Schnell R, Hill PB, Esser E. *Methoden der empirischen Sozialforschung*. München, Wien, Oldenbourg Wissenschaftsverlag, 1995.

- 48 Forsyth BH, Lessler JT. Measurement Errors in Surveys Cognitive Laboratory Methods: A Taxonomy. New York: Wiley, 1991.
- 49 Radtke FM, Franck M, Oppermann S, et al. [The Intensive Care Delirium Screening Checklist (ICDSC)--translation and validation of intensive care delirium checklist in accordance with guidelines]. *Anesthesiol Intensivmed Notfallmed Schmerzther* 2009;44:80-86.
- 50 Bergeron N, Dubois MJ, Dumont M, Dial S, Skrobik Y. Intensive Care Delirium Screening Checklist: evaluation of a new screening tool. *Intensive Care Med* 2001;27:859-864.
- 51 Tavares MG, Pizzichini MM, Steidle LJ, et al. The Asthma Control Scoring System: translation and cross-cultural adaptation for use in Brazil. *J Bras Pneumol* 2010;36:683-692.
- 52 LeBlanc A, Robichaud P, Lacasse Y, Boulet LP. Quantification of asthma control: validation of the Asthma Control Scoring System. *Allergy* 2007;62:120-125.
- 53 Scott J, Huskisson EC. Accuracy of subjective measurements made with or without previous scores: an important source of error in serial measurement of subjective states. *Ann Rheum Dis* 1979;38:558-559.
- 54 Moher D, Schulz KF, Altman DG. [Das CONSORT Statement: Überarbeitete Empfehlung zur Qualitätsverbesserung von Report randomisierte Studien im Parallel-Design.]. *Dtsch Med Wochenschr* 2004;129:T 16-T 20.
- 55 Tackmann R, Schuetz G, Hamm B, Dewey M. [Quality of the reporting of diagnostic accuracy studies: STARD (Standards for the Reporting of Diagnostic accuracy studies)]. *Rofo* 2010;182(8):655-659.
- 56 Ziegler A, König IR. [Guidelines for research reports: German translation of CONSORT 2010, PRISMA and STARD]. *Dtsch Med Wochenschr* 2011;136(8):357-358.

8. Lebenslauf

„Mein Lebenslauf wird aus datenschutzrechtlichen Gründen in der elektronischen Version meiner Arbeit nicht veröffentlicht.“

9. Danksagung

Nach dem Abschluss der Arbeit sowie Fertigstellung der Dissertation ist es mir ein besonderes Bedürfnis, mich bei den nachfolgend aufgeführten Personen zu bedanken.

Ein erster und besonderer Dank geht an dieser Stelle an meine Doktormutter, Frau Universitätsprofessorin Dr. med. Claudia Spies, Klinikdirektorin der Klinik für Anästhesiologie mit Schwerpunkt operative Intensivmedizin, Campus Virchow-Klinikum und Campus Charité Mitte, CHARITÉ - Universitätsmedizin Berlin. Das gilt sowohl für die Überlassung des Dissertationsthemas, die uneingeschränkt engagierte Betreuung und gute Beratung als auch für die kritische Durchsicht des Manuskripts.

Ferner mein allerherzlichster Dank für die stets motivierende und unermüdliche Betreuung und Beratung bei der Durchführung meiner Arbeit an Herrn Dr. Jan-Philipp Breuer, wissenschaftlicher Mitarbeiter der Klinik für Anästhesiologie mit Schwerpunkt operative Intensivmedizin, Campus Virchow-Klinikum und Campus Charité Mitte, CHARITÉ - Universitätsmedizin Berlin.

Nicht zuletzt möchte ich mich bei Prof. Dr. Patrick M. Bossuyt, Department of Clinical Epidemiology and Biostatistics, Academic Medical Center, University of Amsterdam für die Erlaubnis zur Übersetzung der STARD-Kriterien, der prompten Durchsicht der Übersetzung und den Anregungen hierzu ganz herzlich bedanken.

10. An Eides Statt

Hiermit erkläre ich eidesstattlich, dass die Dissertation von mir, Michael Barz, ohne die (unzulässige) Hilfe Dritter verfasst wurde, auch in Teilen keine Kopie anderer Arbeiten darstellt und die benutzten Hilfsmittel sowie die Literatur vollständig angegeben sind.

Berlin, den 27. April 2012