

SEMANTIC AND CULTURAL EQUIVALENCE OF THE WORKING ALLIANCE INVENTORY SHORT-REVISED SCALE FOR THERAPEUTIC ALLIANCE IN FAMILY MEDICINE: LESSONS LEARNED IN SLOVENIA

DOSEGANJE KULTURNE IN JEZIKOVNE USTREZNOSTI LESTVICE ZA OCENJEVANJE TERAPEVTSKEGA ZAVEZNIŠTVA MED ZDRAVNIKOM IN BOLNIKOM V DRUŽINSKI MEDICINI - SLOVENSKE IZKUŠNJE

Davorina PETEK^{1*}, Ambrož PUŠNIK², Polona SELIČ¹, Eva CEDILNIK-GORUP¹,
Žan TRONTELJ^{1,3}, Marine RIOU⁴, Jean Yves LE RESTE⁵

¹University of Ljubljana, Faculty of Medicine, Department of Family Medicine, Poljanski nasip 58, 1000 Ljubljana, Slovenia

²University of Ljubljana, Faculty of Medicine, Vrazov trg 2, 1000 Ljubljana, Slovenia

³Community Health Centre, Derčeva 5, 1000 Ljubljana, Slovenia

⁴Universite de Brest - Bretagne occidentale, Faculté de Médecine & des Sciences de la Santé

⁵EA 7479 SPURBO, Faculté de Médecine et des Sciences de la Santé, 22 Avenue Camille Desmoulins, CS 93837, 29238 BREST CEDEX 3 C

Received: Aug 9, 2018

Accepted: Nov 26, 2018

Original scientific article

ABSTRACT

Keywords:
family medicine,
therapeutic
alliance, Delphi
consensus
procedure

Introduction: Therapeutic alliance is a term most commonly associated with psychotherapeutic treatment, but recently its use has become increasingly significant in the other fields of medicine. An increasing amount of evidence implies that the quality of the therapeutic alliance between the doctor and patient substantially affects treatment outcomes. A European consensus chose the Working Alliance Inventory - Short Revised (WAI-SR) scale as the most efficient for European primary care. This paper presents the process of establishing the semantic and cultural equivalence of the two WAI-SR scales in Slovene.

Method: As a part of a larger international project, a group of four experts translated the two WAI SR scales (physician and patient versions) from English into Slovene. Twenty-six Slovenian family medicine doctors participated in the process of obtaining semantic, idiomatic, experiential and conceptual equivalence in translation using a Delphi consensus procedure. Afterward, a cultural equivalence was made to adapt the translations within the national context.

Results: Agreement on translation was achieved after two Delphi rounds. The back-translation and cultural equivalence were accomplished without major problems, with some minor additional linguistic corrections.

Conclusion: A Slovene version of the WAI-SR scale was successfully adapted and is available for further scale validation and research on therapeutic alliance.

IZVLEČEK

Ključne besede:
družinska
medicina,
terapevtsko
zavezništvo,
Delphi metode

Uvod: Terapevtsko zavezništvo je pojem, ki se je tradicionalno in običajno povezoval s psihoterapevtsko obravnavo, v zadnjem času pa je njegova uporaba postala pomembna tudi v medicini. Vse več dokazov namreč kaže, da so potek in izidi zdravljenja ter zadovoljstvo bolnikov povezani s kakovostjo odnosa (terapevtskega zavezništva) med zdravnikom in bolnikom. Kot eno od najučinkovitejših orodij za ocenjevanje terapevtskega zavezništva v primarni zdravstveni oskrbi je bila v Evropi z mednarodnim konsenzom izbrana lestvica za ocenjevanje terapevtskega zavezništva med zdravnikom in bolnikom (LTZ, angl. WAI-SR). Lestvica služi ocenjevanju odnosa med zdravnikom in bolnikom z obeh strani, tako zdravnika kot bolnika. Namen dela je bil doseči jezikovno in kulturno ustreznost obeh prevedenih lestvic WAI-SR v slovenščini.

Metode: V okviru večjega mednarodnega projekta je skupina štirih slovenskih strokovnjakov prevedla zdravnikovo in bolnikovo različico lestvice WAI-SR iz angleščine v slovenščino. Skupina je bila sestavljena iz dveh zdravnikov družinske medicine, slovenista in anglista. V postopku doseganja semantične, idiomatske, izkustvene in konceptualne ustreznosti prevoda s pomočjo Delphi metode je sodelovalo 26 slovenskih zdravnikov družinske medicine. Sledila je prilagoditev prevodov za zagotavljanje jezikovne in kulturne ustreznosti lestvice v slovenskem govornem in kulturnem okolju.

Rezultati: Soglasje s prevodom je bilo v skupini 26 sodelujočih zdravnikov družinske medicine doseženo po dveh Delphi krogih. Pretvorba slovenskega prevoda lestvic nazaj v angleški jezik, primerjava osnovnega angleškega besedila in povratnega prevoda v izvorni jezik ter preverjanje kulturne ustreznosti so bili doseženi brez večjih težav, z nekaj manjšimi dodatnimi jezikovnimi popravki.

Zaključek: Predstavljeni proces kaže na pomen iskanja jezikovne in kulturne ustreznosti ocenjevalnih lestvic, ki jih v medicini avtorji marsikdaj zgolj slovenijo ter uporabljajo brez preverjanja ustreznosti - ne samo jezikovne in kulturne, ampak tudi v smislu merskih značilnosti. Izkušnje procesa opozarjajo na potrebno skrbnost in zadržanost pri nekritični rabi najrazličnejših lestvic v medicini, ne da bi avtorji preverili, kaj dejansko ocenjujejo z njimi. Slovenska verzija lestvice WAI-SR je bila uspešno prilagojena in je na voljo tako za nadaljnjo validacijo (na primer konstrukta) kot za raziskave terapevtskega zavezništva.

*Corresponding author: Tel. + 386 1 436 82 17; E-mail: davorina.petek@gmail.com

1 INTRODUCTION

One of the key characteristics of family medicine is a long-term doctor-patient relationship (1). Research shows that most patients decide to stay with their personal physician, i.e. family medicine doctor, for at least eight years or even longer (2). A personal relationship thus develops between the patient and the physician, and the nature of this relationship has significant effects on prognosis and patient satisfaction (3).

1.1 The Long-Term Doctor-Patient Relationship

The long-term doctor-patient relationship can also be described by the concept of therapeutic alliance. The most commonly-cited definition of this was first articulated by Bordin (4) in 1979; he argued that the construct consists of three components: the bond between the therapist and the patient; therapist-patient agreement on the goals of treatment; and therapist-patient agreement on the tasks of treatment. Therapeutic alliance also involves an assessment of doctor-patient trust, communication, and patient cooperation (5). It thus exceeds the paternalistic doctor-patient relationship and implements a model of shared decision-making, where the relationship is based on mutual trust, understanding, and the doctor's empathy. Scales for assessing therapeutic alliance were first developed and validated in psychotherapy (6), and proved to be a useful tool in assessing psychotherapeutic alliance when dealing with both in- and outpatients (7). The scales have not yet been specifically altered for use outside psychotherapy, but can nonetheless be helpful in the evaluation and improvement of the therapeutic relationship elsewhere (8). Several research studies show that therapeutic alliance is also associated with better treatment results in clinical medicine (9). A study evaluating patient outcomes in cardiac rehabilitation programmes showed that a strong therapeutic alliance could play an important role in achieving favourable results (10). Its use, therefore, also seems to be applicable to family medicine.

1.2 The Assessment of a Therapeutic Alliance

The importance of using validated assessment tools has been receiving growing attention, with researchers becoming more aware that tools and techniques with established validity and reliability produce more consistent and accurate results.

Internationally, several tools have been used to assess therapeutic alliance in previous research (8, 11-15). In Slovenia, different studies have been conducted assessing various doctor-patient relationship attributes, such as empathy (16) or patient satisfaction (2), but no study has

validated or even used the therapeutic alliance scales. This survey is the Slovenian part of an international research study aiming to validate the WAI-SR scale for therapeutic alliance for patients and physicians throughout Europe (8).

1.3 Validation of the Slovenian WAI-SR Scale

A WAI-SR instrument that has been previously validated in one language is not automatically equivalent to the same instrument in another language and/or culture. The equivalence between translated versions of the questionnaire is important for its international comparison.

Slovene is a language spoken by only about two million people. However, scales from other languages still need to be translated and made equivalent to the original language in terms of concepts (the concept must exist in different cultures (17) and semantics, i.e. "equivalence in meaning between the source and the question wording" (18). International research also requires scales to be culturally equivalent, to enable understanding, interpretation and assessment of the subject, that is equal or similar across different cultures. However, cultural factors cannot be seen at the level of the form or meaning of language, and exist only in the background. Since cultural factors are those relating to value systems, geographical situation, traditions, religion, etc., it is important to consider any impact that a culture or way of life can have on wording. The procedure for testing cultural linguistic equivalence consists of the evaluation of the back-translated version, test-retest by bilingual respondents, adaptation of the translated version, and a final cultural check by a principal researcher in the target country (19).

The aim of this study was to obtain a culturally consistent translation of the two WAI-SR scales (one concerning physicians and one concerning patients) and their scoring key. The semantic and cultural equivalence process is also presented.

2 METHOD

2.1 Design of the Study

An international group of researchers was formed under the umbrella of the European General Practice Research Network (EGPRN), led by the University of Brest, consisting of ten national research teams simultaneously working on a translation procedure following the same protocol, aiming to develop a tool available and equivalent in different languages and cultures. In this paper, we present the Slovenian part of the study.

With the aim of obtaining semantic, idiomatic, experiential and conceptual equivalence in translation, both scales and the scoring key were translated by an e-mail with a forward and back translation using a Delphi procedure. Afterward, a cultural equivalence was performed to adapt the translations within the national context, in order to ensure the homogeneity of the scales throughout Europe.

2.2 Participants

A small group of four experts was formed for the forward translation from English to Slovene.

A convenient sample of 30 practising and academic Family Medicine Doctors (FMDs) were invited to participate in the Delphi method to achieve consensus (20). All participants were provided with a written explanation of the aims and procedure of the study, and signed a statement on voluntary participation. Among those thirty invited experts, four did not accept the invitation, with one saying he was too busy, and three not replying.

Two independent English language translators undertook back translation.

2.3 Instrument

The WAI-SR scales are a 12-item questionnaire for the patient and a 10-item questionnaire for the physician, assessing the three main features of the therapeutic alliance between them: goal, tasks, and bond (8). In this scale, the patient or physician rates each item on a 5-point Likert scale from "1 - rarely or never" and "5 - always". The higher the score, the better the therapeutic alliance. The scoring key provides instructions for the evaluation of the scales.

2.4 Procedure

2.4.1 Forward Translation

The group of four experts made a forward translation from English to Slovene of both WAI-SR scales. All the differences in translation were reconciled between them until they reached a consensus.

2.4.2 Validation of the Forward Translation by the Delphi Method

To verify semantic equivalence, both scales were sent to the participants in the Delphi method, which allows a group to elicit judgments through an iterative process, interspersed with controlled feedback of opinions (20). The group evaluated the translation for clarity, common language, and conceptual adequacy. The experts were contacted separately by email to ensure the anonymity and independence of each opinion. Each participant

was asked to validate or reject the translation by rating each statement on a scale from 1 to 9, where 1 meant "no agreement" and 9 meant "full agreement". If they rated a translation with less than 7, they were asked to explain their disagreement and possibly propose a more suitable translation. The principal researcher evaluated the answers.

A successful validation for each statement was obtained when at least 70% of the participants rated it 7 or above. If a statement did not meet this criterion, the principal researcher proposed a new translation taking into account the participants' suggestions. The new translation was again sent to the group for a second Delphi round. The process was repeated until all the statements were successfully validated.

2.4.3 Back Translation

Two licensed translators with no knowledge of the original English version of the WAI-SR scales independently translated the validated items back from Slovene into English. After the independent translation, they were asked to reach a consensus on the translated items. In the event of disagreement, the leader of the Slovenian research team led the consensus procedure until it was achieved for all statements.

2.4.4 Cultural Adaptation

Translation issues were discussed by the research group, which met twice a year for two years. The international collaborative group compared the back-translation to the original English version at a workshop during an EGPRN meeting in Dublin 2017. The team leaders of five countries and an international committee with the principal investigator of the TATA group carried out a cultural check by comparing a back-translation of five languages, including Slovenian, with the original version. The main task was to identify those translated items, whose meaning might have been lost or inappropriately altered in translation. If the problem could not be solved, it was submitted to the local research team to propose a solution.

The whole process is shown in Figures 1 and 2.

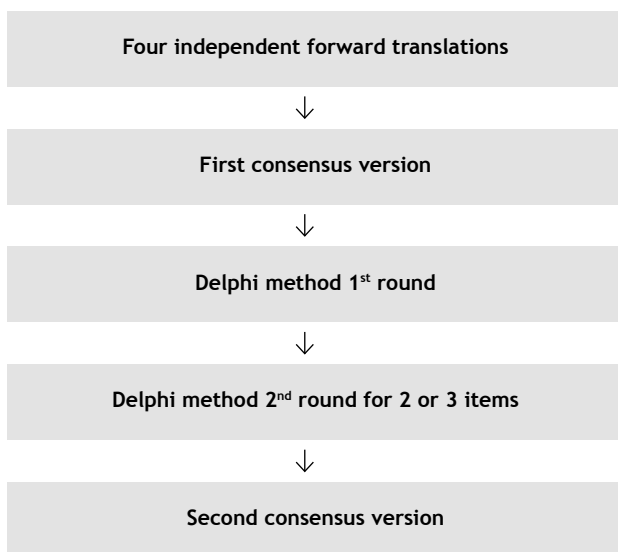


Figure 1. Flowchart of cultural equivalence for WAI SR scales and scoring key.

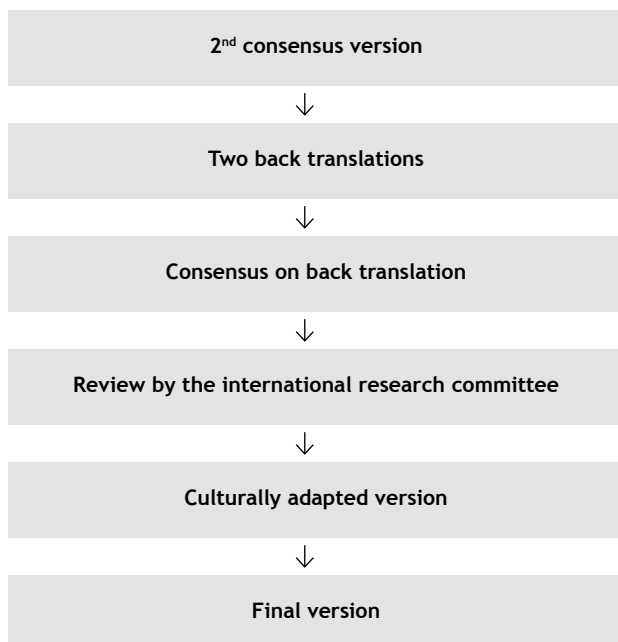


Figure 2. Flowchart of cultural adaptation for WAI-SR scales and scoring key.

3 RESULTS

3.1 The Sample

The group of four translators who did the forward translation consisted of two FMDs, one linguist, and one psychologist; three women and one man, with a mean age of 53 years (range 43-60). All were fluent in English.

26 experts (FMDs) participated in the Delphi part of the study. Among them, 3/26 were male and 23/26 female; their mean age was 40.7 years (range 27-60) and the mean number of years of working in the practice was 12.3 years (range 1-34). 18/26 experts were involved in teaching, and 19/26 were researchers. Of the whole group, 7/26 participants worked in a solo practice and 17/26 in a group practice, while 2/26 were still trainees; 13/26 worked in a rural or semirural environment.

3.2 WAI-SR Patient Scale

For validation, each item had to be rated at 7 or more by at least 17 participants. The first Delphi round for the WAI-SR patient scale showed acceptable agreement in all but one statement (Q8), where only 12/26 (46%) of the participants rated the translation with 7 or above (Table 1).

Table 1. WAI-SR patient scale Likert scores, mean and median - Round 1 (N=26).

Results	Inst	Ans	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12
≥7 (n/26)	23	23	22	23	20	24	23	20	24	12	20	19	20	25
≥7 (%)	88	88	85	88	77	92	88	77	92	46	46	73	77	96
Mean	8.0	8.3	7.7	8.2	7.8	8.6	8.6	7.9	8.7	6.1	8.0	8.0	7.9	8.7
Median	8	9	8	9	9	9	9	8	9	6	9	9	8	9

Legend: N - Number of participants that rated the item ≥7 out of all 26 participants; Inst - instructions; Ans - Likert scale answers; Q - question

In the second Delphi round concerning Q8, the participants proposed 18 alternative translations. We present some of the suggestions in Supplementary material - Table 1. Agreement was reached on a revised translation.

3.3 WAI-SR Physician Scale

The first Delphi round for the WAI-SR Family physician scale showed agreement in all but two statements (Q2, Q10). Q2 was rated as adequate by 15/26 (58%) of participants, but Q10 by only 10/26 (38%) participants (Table 2).

Table 2. WAI-SR physician scale Likert scores, mean and median - Round 1 (N=26).

Results	Inst	Ans	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
≥7 (n/26)	24	25	23	15	20	21	25	23	24	22	22	10
≥7 (%)	92	96	88	58	77	81	96	88	92	85	85	38
Mean	8.2	8.5	8.4	6.7	8.0	7.5	9.0	8.5	9.0	9.0	9.0	6.0
Median	8	9	9	8	9	7.5	9	8.5	9	9	9	6

Legend: N - Number of participants that rated the item ≥7 out of all 26 participants; Inst - instructions; Ans - Likert scale answers; Q - question

Numerous alternative translations were again proposed in the second Delphi round - 6 for Q2 and 5 for Q10. Some of these are presented in Supplementary material - Tables 2 and 3. Again, agreement was reached on a revised translation.

The Slovene version of the translation for both scales was issued after the second Delphi round (Table 3, Supplementary material - Table 4).

Table 3. Mean and median: Patient scale Q8 and Physician scale Q2, Q10 - Round 2.

	Q8 patient scale	Q2 physician scale	Q10 physician scale
Mean	8.2	8.7	8.2
Median	9	9	8.5

3.4 Back-Translation

Consensus was achieved in two rounds of the agreement process between the two professional translators for both WAI SR scales. Consensus on the back-translation of the scoring key was achieved in four rounds.

3.5 International Cultural Equivalence Evaluation

The work on cultural equivalence highlighted three potential problems with translation: In Q10 of the patient scale, "treatment" was translated as "consultation". We concluded that, considering its original use in psychotherapy, the word "treatment" did not relate solely to medical treatment but to the entire process of doctor-patient consultation. For this reason, the translation relates to the entire process of consultation and not only to treatment actions.

Also, in the patient scale, Q12 was originally in the active voice, but was translated to the passive voice in the validated Slovene translation. After the cultural check was carried out, the national team agreed that the use of the active voice was more suitable, since it emphasized the patient's active role in the consultation and corresponded to the "shared decision-making" model (Supplementary material - Table 5).

In Q8 of the physician scale, the discussion on cultural equivalence revealed that there was a difference between 'the common perception of a goal' (as in the original version) and 'common agreement on the goal' (as translated). We consulted the linguist and appropriately altered the validated translation so that the original meaning of the statement was retained (Supplementary material - Table 6).

The final version of the WAI-SR Slovene translation was accepted after the second Delphi round, including these cultural adaptations.

3.6 Validation of the Scoring Key

The scoring key contains instructions concerning the evaluation of the scale. The same procedure was used for the translation of the scoring key as for the WAI-SR items. It was validated in the first Delphi round with Q1, Q2, and Q3 each having one evaluation of <7, and all the others (except the last item) >7. Item 11 was adapted after consulting the author of the scale, AO Horvath, who gave additional instructions.

4 DISCUSSION

4.1 Main Findings

Only two rounds of the Delphi method were needed to achieve a consensus on the translations of all the items. Cultural equivalence of the back-translation was obtained after some minor adaptations were made. The process showed that a simple literal translation was inappropriate, and rigorous efforts must be made to ensure the meaning and intent of the original items are maintained so the scale remains relevant.

4.2 Validation Process and Comparison to other Countries

The equivalence procedure in the translation of the two scales assessing therapeutic alliance was complex and time-consuming, but it served well for the purpose of semantic validation. The same procedure was used to validate and achieve equivalence in the translation of the definition of multi-morbidity (21, 22) and to validate the WAI-SR questionnaire in other countries (23). The translation into Polish showed the feasibility of the procedure, taking only one Delphi round to achieve consensus (23). The advantage of this procedure also lies in the fact that it was simultaneously taking place in several European countries with different linguistic bases, which provided the opportunity to discuss the difficulties national and local research groups met with while translating the original WAI-SR scales.

The Delphi method was used to validate the agreed forward translation and has been shown to be suitable for exploring areas where controversy, debate or a lack of clarity exist. Within this process, translations of WAI-SR scales were actively tested in representatives of the target population or language group to determine whether the respondents understood the questionnaire in the same way as the original. We feel that the use of this method for translation was legitimate, since it provided an accurate consensus technique (24).

Ideally, every questionnaire translation should undergo a cultural equivalence to identify and resolve any inadequate expressions in the translation, as well as to sort out any other discrepancies between the original items and the back-translated ones. The first steps in the process were inspired by the work of Streiner et al. (25). The standardized approach for the cultural adaptation of patient-measured outcomes was confirmed in recent guidelines (26, 27). In this study, we followed the recommendations at all stages: in the first part of cultural adaptation by using the Delphi method, because we recognized this as the best option given the specifics of our language, social and cultural context, and then by the supervision of the researchers led by the University of Brest, who oversaw the adaptation of the questionnaire

and the cultural adaptation based on the back-translation. This was to ensure that the items were translated considering their structure as well as the suitability of their content.

4.3 Limitations

Given that translation is the most common method for preparing instruments for cross-cultural research, we must be alert to the pitfalls that threaten validity.

Firstly, when translating scales such as the WAI-SR, it would be best if the forward translation was carried out by professionals who fit these criteria: familiar with the terminology used in the questionnaire; knowledgeable about the subjects covered; experienced in translating scales from (as here) English; and have Slovene as their native language. The content of the WAI-SR covers the fields of psychology and medicine, and its translation must be understandable by both physicians and patients. In Slovenia, we were unlikely to find a professional translator who would meet all these criteria. Creating a group of two family medicine doctors, a psychologist and a linguist to carry out the forward translation solved this problem.

Secondly, we stated that the experts carrying out the consensus procedure consisted of individuals who were fluent in English. However, the method of evaluating fluency in the language is debatable. Proficiency in English was assessed in two ways: one was self-evaluation, and the other was the number of English publications of each of the participants. The latter, in particular, may not be a powerful tool for showing language fluency; however, it was a pragmatic and feasible solution.

Thirdly, the Delphi group was not representative of the community of Slovenian family medicine doctors (FMDs) - men were underrepresented and the percentage of the academic FMDs involved was higher than the Slovenian average. But considering that the Delphi method is a qualitative one, population representativeness is not necessary. It is more important that all the characteristics of the participants that can influence decisions regarding validation are represented, such as different ages, location of practice, years of experience and involvement in the academic side of family medicine.

Finally, it would have been preferable if the back-translation had been made by an independent translator fluent in Slovene but whose native language was English. Since no such translators were available, we settled for two independent licensed Slovenian translators who had no previous knowledge of the WAI-SR scale.

5 CONCLUSION

At this stage, the WAI-SR and its use in family medicine generally lacks a theoretical background that needs to be discussed and agreed upon in the broader field of family medicine. Given the complexities of patient care in family medicine, the question arises as to whether therapeutic alliance is relatively stable over the course of a relationship between a family doctor and a patient. In addition, if assessing the alliance at one or several points in time, alliance ratings are expected to be associated with morbidity changes over the course of a patient's life, which may fail to capture the short-term impact of alliance on a specific symptom or improvement in their condition. Therefore, the future accuracy of ratings provided by this instrument can be affected by many methodological factors, including the quality of the instrument in terms of validity, reliability, and sensitivity to change. We only described the first phase, where the scale's semantic and cultural equivalence were verified. Further studies will provide results of reliability and item validity analyses. Exploratory principal component analyses are to be conducted to compare response patterns with the hypothesized scale constructs. Four major issues need to be considered in the future: the psychometric properties of the Slovene WAI-SR scale; the appropriateness of the scale for FMDs; practical aspects of scale administration; and the theoretical foundation of scale interpretation within the field of family medicine.

CONFLICT OF INTERESTS

The authors declare that they have no competing interests.

FUNDING

The Slovenian part of the study was partly funded by the Slovenian Institute for the Development of Family Medicine. The co-author PS acknowledges financial support from the Slovenian Research Agency, research core funding Research in the Field of Public Health No. P3-0339. The international TATA study was funded by the European General Practice Research Network.

ETHICAL APPROVAL

The study was approved by the Republic of Slovenia National Medical Ethics Committee at the Ministry of Health, on Dec 15th, 2017, number 0120-397-2016/2.

ACKNOWLEDGMENTS

We thank all the physicians who participated in the study and Ms. Justi Carey for the language editing.

We thank to TATA group researchers, who participated in the international translations and are non-author collaborators: Nicola Buono (National Society of Medical Education in General Practice, Caserta, Italy), Radost Assenova (University of Plovdiv, Bulgaria), Krzysztof Buczkowski (Nicolaus Copernicus University in Torun, Poland), Ana Claveria (Universidad de Vigo, Spain); Robert Hoffman (Tel Aviv University, Israel), Djurdjica Lazic (University of Zagreb, Croatia), Heidrun Linger (Hannover Medical School, Germany), Hans Thulesius (Lund University, Sweden).

ABBREVIATIONS

WAI-SR scale: the Working Alliance Inventory - Short Revised scale

FMDs: Family medicine doctors

REFERENCES

- Tammes P, Salisbury C. Continuity of primary care matters and should be protected. *BMJ*. 2017;356:j373. doi: 10.1136/bmj.j373.
- Petek D, Künzi B, Kersnik J, Szecsenyi J, Wensing M. Patients' evaluations of European general practice - revisited after 11 years. *Int J Qual Health Care*. 2011;23:621-8. doi: 10.1093/intqhc/mzr052.
- Hall MA, Dugan E, Zheng B, Mishra AK. Trust in physicians and medical institutions: what is it, can it be measured, and does it matter? *Milbank Q*. 2001;79:613-39. doi: 10.1111/1468-0009.00223.
- Bordin ES. The generalizability of the psychoanalytic concept of the working alliance. *Psychother Theory Res Pract*. 1979;16:252-60. doi: 10.1037/h0085885.
- Huff NG, Nadig N, Ford DW, Cox CE. Therapeutic alliance between the caregivers of critical illness survivors and intensive care unit clinicians. *Ann Am Thorac Soc*. 2015;12:1646-53. doi: 10.1513/AnnalsATS.201507-4080C.
- McLeod BD. Relation of the alliance with outcomes in youth psychotherapy: a meta-analysis. *Clin Psychol Rev*. 2011;31:603-16. doi: 10.1016/j.cpr.2011.02.003.
- Munder T, Wilmers F, Leonhart R, Linster HW, Barth J. Working Alliance Inventory-Short Revised (WAI-SR): psychometric properties in outpatients and inpatients. *Clin Psychol Psychother*. 2010;17:231-9.
- Paap D, Dijkstra PU. Working Alliance Inventory - Short form Revised. *J Physiother*. 2017;63:118. doi: 10.1016/j.jphys.2017.01.001.
- Ferreira PH, Ferreira ML, Maher CG, Refshauge KM, Latimer J, Adams RD. The therapeutic alliance between clinicians and patients predicts outcome in chronic low back pain. *Phys Ther*. 2013;93:470-8. doi: 10.2522/ptj.20120137.
- Burns JW, Evon D. Common and specific process factors in cardiac rehabilitation: independent and interactive effects of the working alliance and self-efficacy. *Health Psychol*. 2007; 26:684-92.
- Luborsky L, Barber JP, Siqueland L, Johnson S, Najavits LM, Frank A, et al The Revised Helping Alliance Questionnaire (HAQ-II): psychometric properties. *J Psychother Pract Res*. 1996;5:260-71.
- Delsignore A, Rufer M, Moergeli H, Emmerich J, Schlesinger J, Milos G, et al. California Psychotherapy Alliance Scale (CALPAS): psychometric properties of the German version for group and individual therapy patients. *Compr Psychiatry*. 2014;55:736-42. doi: 10.1016/j.comppsy.2013.11.020.
- Kim SC, Boren D, Solem SL. The Kim Alliance Scale: development and preliminary testing. *Clin Nurs Res*. 2001;10:314-31.
- Shelef K, Diamond GM. Short form of the revised Vanderbilt therapeutic alliance scale: development, reliability, and validity. *Psychother Res*. 2008;18:33-43. doi: 10.1080/10503300701810801.
- Saunders SM, Howard KI, Orlinsky DE. The Therapeutic bond scales: psychometric characteristics and relationship to treatment effectiveness. *Psychol Assessment*. 1989;1:323-30.
- Pešek L, Selič P. Empathy and burnout in Slovenian family medicine doctors: the first presentation of Jefferson Scale of Empathy results. *Zdr Varst*. 2018;57:155-65. doi: 10.2478/sjph-2018-0020.
- Patrick DL, Wild DJ, Johnson ES, Wagner TH, Martin MA. Cross-cultural validation of quality of life measures. In: Orley J, Kuyken W, editors. *Quality of life assessment: international perspectives*. Heidelberg: Springer Verlag, 1994:19-32.
- Acquadro C, Jambon B, Ellis D, Marquis P. Language and translation issues. In: Spilker B, editor. *Quality of life and pharmacoeconomics in clinical trials*. 2nd ed. Philadelphia: Lippincott-Raven, 1996:575-85.
- Herdman M, Fox-Rushby J, Badia X. "Equivalence" and the translation and adaptation of health-related quality of life questionnaires. *Qual Life Res*. 1997;6:237-47.
- Skulmoski GJ, Hartman FT, Krahn J. The Delphi method for graduate research. *J Inf Technol Educ*. 2007;6:2-21.
- Lazić ĐK, Le Reste JY, Murgić L, Petriček G, Katič M, Ožvačić-Adžić Z, et al. Say it in Croatian - Croatian translation of the EGPRN definition of multimorbidity using a Delphi consensus technique. *Coll Antropol*. 2014;38:1027-32.
- Le Reste JY, Nabbe P, Rivet C, Lygidakis C, Doerr C, Czachowski S, et al. The European general practice research network presents the translations of its comprehensive definition of multimorbidity in family medicine in ten European languages. *PLOS One*. 2015;10:e0115796. doi:10.1371/journal.pone.0115796.
- Le Reste JY, Buczkowski K, Morvan F, Lazic V, Lingner H, Petek D, et al. Translation of a therapeutic alliance scale (The WAI SR) into Polish, TATA EGPRN collaborative study. Abstracts/programme book of the 85th EGPRN meeting, Dublin, 19-22 October 2017. Accessed March 20th, at: <https://www.egprn.org/file/3898d1a2-a2f8-479a-899c-c90a1266e949/2017-oct-dublin-ireland-programme.pdf>.
- Linstone HA, Turoff M, editors. *The Delphi method: techniques and applications*. 2002. Accessed July 17th, 2018 at: <http://is.njit.edu/pubs/delphibook>.
- Streiner DL, Norman GR, Cairney J. *Health measurement scales: a practical guide to their development and use*. 5th ed. Oxford, New York: Oxford University Press, 2014.
- Wild D, Eremenco S, Mear I, Martin M, Houchin C, Gawlicki M, 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;12:430-40. doi: 10.1111/j.1524-4733.2008.00471.x.
- Rabin R, Gudex C, Selai C, Herdman M. From translation to version management: a history and review of methods for the cultural adaptation of the EuroQol five-dimensional questionnaire. *Value Health*. 2014;17:70-6. doi: 10.1016/j.jval.2013.10.006.

SUPPLEMENTARY MATERIAL

Supplementary table 1. Q8 patient scale translation examples.

Q8 patient scale - original statement	Q8 patient scale - forward translation	Q8 patient scale - alternatives
_____ and I agree on what is important for me to work on.	Z _____ se skupaj dogovarjava o tem, kaj je zame pomembno, da počnem.	Z _____ se strinja, na čem moram delati Z _____ se strinja/soglašava o tem, kaj je zame pomembno, da počnem. Z _____ se strinja, kaj je zame pomembno, da izboljšam.

Supplementary table 2. Q2 physician scale translation examples.

Q2 patient scale - original statement	Q2 patient scale - forward translation	Q2 patient scale - alternatives
I am genuinely concerned for _____'s welfare.	Blagostanje _____ je moja osrednja skrb.	Skrbi me pacientovo dobro. Dobrobit _____ je moja osrednja skrb. Moja pristna skrb je dobro počutje _____.

Supplementary table 3. Q10 physician scale translation examples.

Q10 patient scale - original statement	Q10 patient scale - forward translation	Q10 patient scale - alternatives
We agree on what is important for _____ to work on.	Z _____ se skupaj dogovarjava, kaj je zanj(o) pomembno, da počne.	Z _____ se strinja, na čem mora delati. Z _____ se strinja o tem, kakšni ukrepi so pomembni. Z _____ se strinja, kaj je zanj(o) pomembno, da izboljša.

Supplementary table 4. validated translations: Q8 patient scale and Q2, Q10 physician scale - Round 2.

Q8 patient scale - successfully validated translation	Q2 physician scale - successfully validated translation	Q10 physician scale - successfully validated translation
Z _____ se strinja, kaj je zame pomembno, da izboljšam.	Dobrobit _____ je moja osrednja skrb.	Z _____ se strinja, kaj je zanj(o) pomembno, da izboljša.

Supplementary table 5. Q12 patient scale cultural equivalence.

Q12 patient scale - original statemet	Q12 patient scale - validated translation	Q12 patient scale - after cultural adaptation
I believe the way we are working with my problem is correct.	Verjamem, da je način obravnave moje težave pravilen.	Verjamem, da moja težavo obravnavava na ustrezen način.

Supplementary table 6. Q8 physician scale cultural equivalence.

Q8 patient scale - original statemet	Q8 patient scale - validated translation	Q8 patient scale - after cultural adaptation
_____ and I have a common perception of his/her goals	Z _____ se strinjava glede njegovih/njenih ciljev.	Z _____ enako dojemava njegove/njene cilje.