



# Challenges in measuring interprofessional–interorganisational collaboration with a questionnaire

Loes J Meijer, Drs, MD <sup>1\*</sup>, Esther de Groot, PhD<sup>2</sup>, Maarten van Smeden, PhD<sup>3</sup>, François G Schellevis, MD, PhD<sup>4,5</sup>, Roger AMJ Damoiseaux, MD<sup>6</sup>

<sup>1</sup>GP and PhD Student, Department of GP Training, Julius Center for Health Sciences and Primary Care, UMC Utrecht, Utrecht, The Netherlands; <sup>2</sup>Assistant Professor, Department of GP Training, Julius Center for Health Sciences and Primary Care, UMC Utrecht, Utrecht, The Netherlands; <sup>3</sup>Statistician and Senior Researcher in Epidemiology and Medical Statistics, Department of Biostatistics and Research Support, Julius Center for Health Sciences and Primary Care, UMC Utrecht, Utrecht, The Netherlands; <sup>4</sup>Professor of Primary Care, Department of Primary Care, NIVEL (Netherlands Institute for Health Services Research), Utrecht, The Netherlands; <sup>5</sup>Professor, Department of General Practice & Elderly Care Medicine, Amsterdam Public Health Research Institute, VU University Medical Centre, Amsterdam, The Netherlands; <sup>6</sup>Professor of Primary Care, Department of GP Training, Julius Center for Health Sciences and Primary Care, UMC Utrecht, Utrecht, The Netherlands

## Abstract

**Background:** Collaboration between medical professionals from separate organisations is necessary to deliver good patient care. This care is influenced by professionals' perceptions about their collaboration. Until now, no instrument to measure such perceptions was available in the Netherlands. A questionnaire developed and validated in Spain was translated to assess perceptions about clinicians' collaboration in primary and secondary care in the Dutch setting.

**Aim:** Validation in the Dutch setting of a Spanish questionnaire that aimed to assess perceptions of clinicians about interorganisational collaboration.

**Design & setting:** After translation, cultural adaptation, and pre-testing, the questionnaire was sent to GPs and secondary care clinicians (SCCs) in three regions in the Netherlands. The responses of 445 responders were used to assess the validity and reliability of the questionnaire.

**Method:** A confirmatory factor analysis (CFA) and an exploratory factor analysis (EFA) were performed to study the construct validity of the hypothesised factor model underlying the questionnaire. Test-retest reliability was evaluated using weighted Kappa statistics.

**Results:** Results of the CFA indicated poor fit of the hypothesised factor structure. EFA, executed separately for each region, showed a highly unstable factor structure. The test-retest reliability analysis demonstrated low re-test reliability.

**Conclusion:** The underlying factor structure of a Spanish questionnaire could not be reproduced. The construct validity and reliability of this questionnaire were insufficient to warrant use in the Dutch setting. This study demonstrates the need for evaluating validity and reliability of questionnaires in local settings.

\*For correspondence: l.j.meijer@umcutrecht.nl

**Competing interests:** The authors declare that no competing interests exist.

**Received:** 04 September 2017

**Accepted:** 14 December 2017

**Published:** 21 February 2018

© This article is Open Access: CC BY license (<https://creativecommons.org/licenses/by/4.0/>)

**Author Keywords:** primary care, secondary care, interprofessional, interdisciplinary collaboration, questionnaire

Copyright © 2018, The Authors;

DOI:10.3399/

bjgpopen18X101385

## How this fits in

A validated Spanish questionnaire was found to be not valid in the Netherlands. When measuring professionals' perceptions about collaboration, one needs to be aware that a questionnaire developed and validated in other settings is not to be used lightly in a different context. Evaluating validity and reliability of such questionnaires in local settings is essential.

## Introduction

Interprofessional collaboration is necessary for patient-centred care, especially in the complex context of an ageing population in which multimorbidity is common.<sup>1</sup> Patients who suffer from multimorbidity often receive care from multiple healthcare professionals, and frequently undergo transitions between healthcare organisations and settings. When professionals from different organisations contribute to patient care, ensuring continuity of care is important.<sup>2-6</sup> The transitions of patients, in particular when moving from the primary care setting to secondary care, benefit significantly from good collaboration between GPs and SCCs, and between the organisations in which they work.<sup>6,7</sup> Even when a healthcare professional functions well on their own, patient care can be suboptimal when professionals work in a fragmented healthcare system with many boundaries between primary care, outpatient settings, and hospitals.<sup>7-14</sup>

In the Netherlands, as in many other countries, various initiatives are undertaken to improve collaborative patient care across organisations.<sup>10,15</sup> Whether these efforts to improve the conditions for collaboration on an individual and an organisational level are successful often remains unclear because adequate instruments to measure collaboration are lacking. Only a few questionnaires have been developed which aim to measure the level of collaboration of professionals in a single organisation or a single team.<sup>16-20</sup> Questionnaires that measure conditions for collaboration across the borders of organisations are scarce, with a few exceptions.<sup>21,22</sup> The Dutch questionnaire developed and studied by Berendsen *et al* quantifies how SCCs and GPs value their mutual collaboration. Nūno-Solinís *et al*'s Spanish questionnaire measures the conditions for collaboration, while taking the influence of the work environment into account.<sup>21,22</sup>

This article describes a study of the validity and test-retest reliability of Nūno-Solinís *et al*'s questionnaire in the Dutch healthcare context. The original questionnaire was developed based on a theoretical model for interprofessional collaboration by D'Amour.<sup>22-25</sup> D'Amour described collaboration as 'the structuring of collective action through sharing of information and decision-making in clinical processes between professionals in different organisations,'<sup>23-24</sup> and distinguished four dimensions of interprofessional collaboration. Two are relational dimensions: shared goals and vision, and internalisation; and two organisational dimensions: governance, and formalisation. Out of four dimensions, 10 interrelating indicators were formulated and evaluated in a multiple-case study<sup>23-24</sup> (see **Box 1**).

Box 1. Dimensions and indicators of a conceptual model for interprofessional collaboration<sup>23-24</sup>

Factors	Dimensions	Indicators
Interpersonal relationships	Shared goals and vision	<ul style="list-style-type: none"> <li>• Shared goals</li> <li>• Patient-centred orientation</li> </ul>
	Internalisation	<ul style="list-style-type: none"> <li>• Mutual acquaintanceship</li> <li>• Trust</li> </ul>
Organisational setting	Governance	<ul style="list-style-type: none"> <li>• Centrality</li> <li>• Leadership</li> <li>• Support for innovation</li> <li>• Connectivity</li> </ul>
	Formalisation	<ul style="list-style-type: none"> <li>• Formalisation tools</li> <li>• Information exchange</li> </ul>

In Spain, these 10 indicators have been used in the development of a questionnaire which measures clinicians' perceptions about conditions for interprofessional collaboration.<sup>22</sup> A study of the validity of the 10-item Spanish questionnaire showed a two-factor structure with promising levels of model fit.<sup>22</sup> The present study aims to re-study the validity of the questionnaire developed in Spain, in order to test its appropriateness for measuring conditions for interprofessional collaboration in the Dutch healthcare setting. Data were collected in three geographic regions in the Netherlands where GPs and SCCs work in a hospital context to deliver collaborative patient care.

## Method

### Content validity and face validity

To test face-validity, 28 professionals were asked by mail to look at the questionnaire, translated from English to Dutch (full questionnaire available in English and Dutch from the authors, on request). These professionals (SCCs  $n = 20$ , GPs  $n = 8$ ) were asked whether they could answer the questions and whether they considered the questions relevant for measuring collaboration in their working situation. Their main comment was about long sentences. In the subsequent translation from Spanish to Dutch, the length of the sentences was a major focus of attention.

The content validity was considered adequate because it was based on the diverse and thorough studies of D'Amour and the questionnaire-building process in Spain, together with the reactions of intended responders.

### Preparation of the questionnaire

Two bilingual translators forward-translated the original Spanish questionnaire into Dutch, and some adjustments were made to the questionnaire to take account of the Dutch healthcare setting. The structure of the 10-item Spanish questionnaire was preserved including, after each question, the phrase 'Please rate the current situation in your organisation with respect to the other level of care on a scale of 1 to 5.' The response options describe desirable attributes with five distinct descriptions for each item on a 5-point Likert scale, all ranging from 1 (none of the attribute) to 5 (lots of the attribute), as demonstrated in the example in **Box 2**. A backward translation from Dutch to Spanish was done. This translation into Spanish was reviewed and approved by one of the developers of the questionnaire.

After translation, the questionnaire was pre-tested, to check whether the translation introduced errors after the test for face validity, using the thinking aloud method.<sup>25-26</sup> Pre-testing was done by five GPs (from city and urbanised rural areas) and five SCCs from different specialties, all native

Box 2. Example of two items of the 10-item questionnaire to assess the interprofessional collaboration of two different levels of care

<b>1. SHARED GOALS</b>				
The existence of explicit shared goals facilitates collaboration and coordination between primary and secondary care settings.				
Please rate the current situation in your organisation with respect to the other level of care on a scale of 1 to 5:				
1. Common goals are missing	2. There are hardly any shared goals	3. There are some common goals	4. There are quite a lot of common goals	5. Nearly all aspects of care are covered by shared goals
<hr/>				
<b>2. PATIENT-CENTRED APPROACH</b>				
When priority is given to the interests and preferences of the patient, this favours collaboration and coordination between professionals working in the primary and secondary care setting.				
Please rate the current situation in your organisation with respect to the other level of care on a scale of 1 to 5:				
1. In the interaction between levels of care, the interests and preferences of patients are not taken into account	2. In the interaction between levels of care, the interests and preferences of patients are taken into account on few occasions	3. In the interaction between levels of care, the interests and preferences of patients are sometimes taken into account	4. In the interaction between levels of care, the interests and preferences of patients are often taken into account	5. In the interaction between levels of care, the interests and preferences of patients are always taken into account

Dutch speakers. Each of them individually read each question aloud.<sup>25–26</sup> A researcher then asked them how they interpreted the questions, and what they thought about the ease of comprehension. None of these professionals indicated relevant aspects of collaboration conditions that were missing in the questionnaire. All comments were noted and discussed within the research team. An example of change of a phrase is ‘professionals working in the different levels’ being changed to ‘the primary and secondary care setting.’

After discussing all the comments in the team, a final version of the questionnaire was established.

### Data collection and sample

For validity testing, an invitation to complete the online questionnaire was sent to all GPs and SCCs ( $N = 1369$ ) practising in one of three selected geographic regions in the Netherlands. Each of these regions contained one large hospital (a top, non-academic, clinical teaching hospital) with a central position in the region. Invitations were sent by e-mail with a link providing direct access to the questionnaire. In two regions, a reminder was sent to all non-responders 2 weeks after the first invitation. In the third region, all responders received the questionnaire again after 2 weeks to evaluate the test-retest reliability. The data were collected between October 2015 and March 2016.

### Data analysis

The approach to studying the (construct) validity of the translated questionnaire was similar to the approach used to study the validity of the original questionnaire.<sup>22</sup> First, the construct validity of the questionnaire was examined using a CFA. Because of the ordinal structure of the data, this analysis was conducted by fitting the factor model on the polychoric correlation matrix<sup>27</sup> of the item responses. The factor model is estimated by maximum likelihood on the data from the three regions separately, to avoid regional effects that may contaminate the correlation analyses and to explore the stability of the factor structure across the regions. Nüno-Solinis *et al.*'s correlated two-factor model was assumed, with the uncorrelated error terms. Initially, measurement invariance across regions was not assumed. The fit of the confirmatory factor models was evaluated using common fit statistics: root mean square error of approximation, standardised root mean square residual, and comparative fit index, with usual cut-off points.

Next, an EFA with oblimin rotation (allowing the factors to be correlated) was employed, because the fit of the CFA models was insufficient. The approach taken was, again, similar to the EFA approach taken by Nüno-Solinis *et al.*<sup>22</sup> The Bayesian Information Criterion (BIC) was used to determine the optimal number of factors.<sup>28</sup> The lower the BIC, the better the fit of the factor model to the data.

Finally, test-retest reliability was evaluated on item-level by squared weighted Kappa (SW Kappa) statistics,<sup>29</sup> taking an SW Kappa of  $>0.70$  as the threshold criterion for reliability.

All analyses were performed in R (version 3.1.1.) using the Lavaan package (CFA), Psych package (EFA), and *irr* package (Kappa).

**Table 1.** Response rate in three Dutch regions

	Total sent	Responses, <i>n</i>	Response rate, %
Region 1	587 (334 GPs, 253 SCCs)	203	35
Region 2	398 (249 GPs, 149 SCCs)	84	21
Region 3	384 (204 GPs, 180 SCCs)	158	41
<b>Total</b>	<b>1369 (787 GPs, 582 SCCs)</b>	<b>445</b>	<b>33</b>

SCC = secondary care clinician.

## Results

### Response

From the 1369 (582 GPs and 787 SCCs) invitees, 458 doctors responded: 206 GPs and 252 SCCs. Thirteen questionnaires were incomplete. There was a total of 445 fully completed questionnaires, representing a response rate of 33% ([Table 1](#)).

### CFAs

The CFA explored the two-factor structure of the questionnaire (interpersonal relationships and organisational setting) in the Dutch healthcare setting, as assumed and tested by Nuño-Solinís *et al.*<sup>22</sup> The fit indices, with their corresponding suggested thresholds levels for sufficient fit to the data, are listed in [Table 2](#). None of the fit indices met the norm criteria for a sufficient fit in each of the three regions.

The questions X1–X4 aimed to measure the factor ‘interpersonal relationships’ and the questions X5–X10 aimed to measure ‘organisational setting’. [Figure 1](#) details the estimated CFA diagram<sup>30</sup> assuming a two-factor structure, the estimated standardised loadings, and corresponding maximum likelihood standard errors. The estimated factor loadings show notable variation between regions and items. High factor loadings (>0.7) were rare ( $n = 4/30$ ). The estimated factor correlations are 0.942 (region 1), 0.754 (region 2), 1.000 (region 3), and 0.923 (total). In all, the estimated factor structures varied considerably between the three regions and, in each of the regions, a lack of support was shown for the assumed two-factor structure. Exploration of modification indices (that is, suggestions for model improvements, such as allowing correlation of error variances) did not yield consistent suggestions for improvements that could improve model fit across the regions (results not shown).

### EFAs

To explore alternative factor structures, EFAs were carried out. EFA models with one, two, three, four, five, and six factors were fitted for each of the regions. Similar to the CFA, the EFA with oblimin rotation showed unstable factor structures across regions. Based on the BIC, a four-factor structure was found to be optimal in region 1; five-factor in region 2; and another five-factor in region 3 (other criteria to determine the optimal number of factors were also checked but did not resolve the instability). The factor loadings for the four and five-factor models were not stable across regions (results not shown). Further explorations with alternative orthogonal rotation strategies and factor selection criteria did not provide a solution to the instability of the factor structure.

### Test-retest reliability

Test-retest reliability analyses were performed with the data gathered in region 3 ( $N = 90$ , GPs = 51, SCCs = 39). The test-retest was measured with the SW Kappa agreement, and is shown in [Table 3](#). The test-retest reliability of this questionnaire was shown to be insufficient for each item (SW Kappa >0.70 is sufficient).

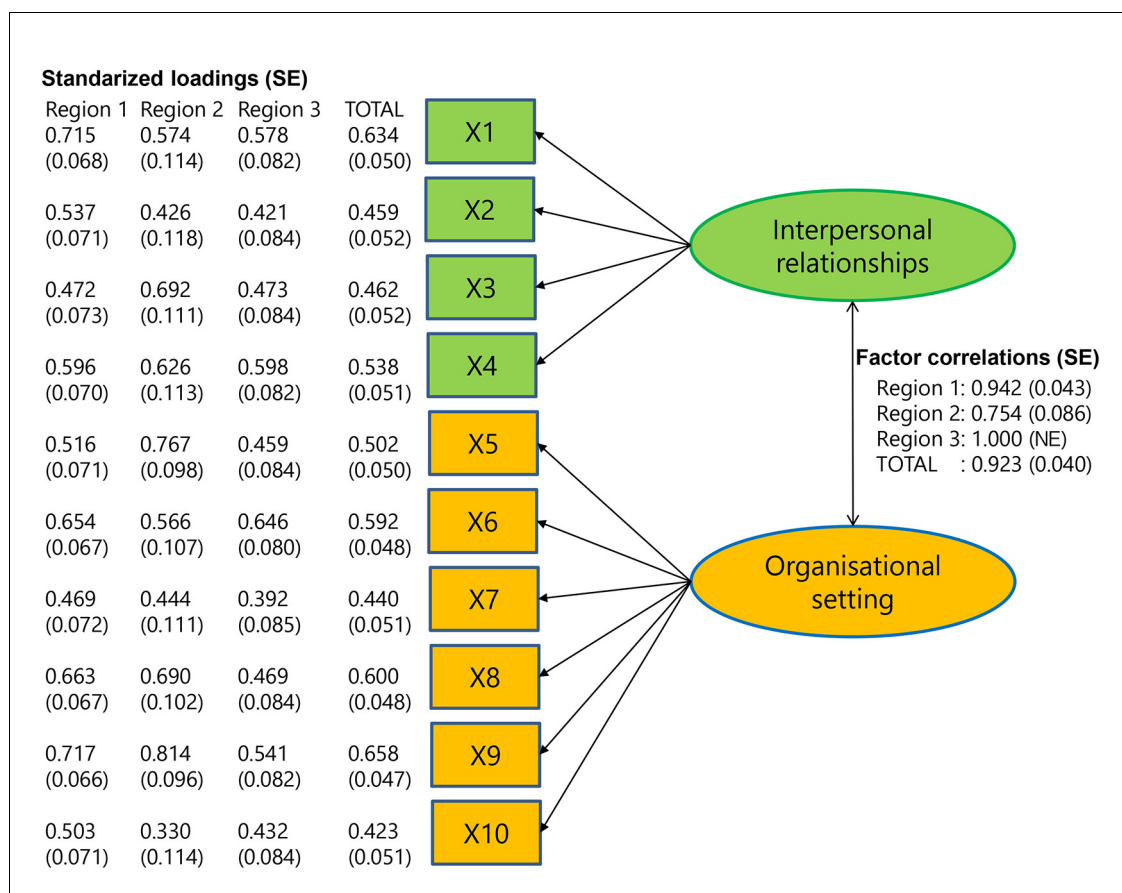
## Discussion

### Summary

In the present study, an evaluation was conducted of the validity and test-retest reliability of a carefully translated and culturally-adapted questionnaire, which aimed to measure interprofessional

**Table 2.** Fit indices for the questionnaire in the Netherlands, by region

Fit indices	Threshold for sufficient fit	Region 1 (n = 203)	Region 2 (n = 84)	Region 3 (n = 158)
Root mean square error of approximation	<0.08	0.14	0.17	0.15
Standardised root mean square residual	<0.08	0.08	0.09	0.09
Comparative fit index	>0.90	0.79	0.72	0.66



**Figure 1** CFA diagram. Results of the confirmatory factor analysis in three regions and the total. NE = not estimable. SE = standard error.

collaboration conditions between clinicians at different levels of care. In contrast to the original questionnaire, the questionnaire showed poor validity in the Dutch setting, and the additional test-retest reliability was also insufficient. The data collected in three different Dutch regions allowed for a cross-regional comparison of validity, by comparing confirmatory factor models with sufficient

**Table 3.** Test-retest reliability, region 3 (N = 90), measured with the SW Kappa<sup>a</sup>. Confidence intervals were estimated by non-parametric bootstrap procedure, based on 5000 bootstrap samples

Questions	Items	SW Kappa	95% CI
X1	Shared goals	0.39	0.27 to 0.50
X2	Patient-centred approach	0.63	0.43 to 0.79
X3	Mutual knowledge	0.42	0.17 to 0.61
X4	Trust	0.43	0.25 to 0.60
X5	Strategic guidelines	0.36	0.22 to 0.48
X6	Shared leadership	0.41	0.20 to 0.58
X7	Support for innovation	0.31	0.07 to 0.52
X8	Forums for meeting	0.56	0.41 to 0.69
X9	Protocolisation	0.32	0.19 to 0.44
X10	Information systems	0.54	0.37 to 0.67

<sup>a</sup>SW Kappa >0.70 the threshold criterion for reliability. SW Kappa = squared weighted Kappa.

sample sizes in each region. This study provides clear evidence that the hypothesised two-factor structure that underlies this questionnaire cannot be confirmed in any of the regions. Consequently, the authors advise against the use of this questionnaire to measure conditions for interprofessional collaboration in the Dutch healthcare context.

One explanation for the difference between validity of the Spanish and Dutch questionnaire may be the differences in organisation of the healthcare system. In the Spanish region where the questionnaire was validated, clinical professionals (doctors and nurses) are part of one integrated care delivery organisation, consisting of a regional hospital (specialised care level), and health centre (primary care level), and are contracted with one provider. In the Netherlands, primary and secondary care takes place in separate organisations and is contracted with different providers. Dutch primary care consists of organisations varying in form and size. Between the Dutch regions, cultural and organisational differences exist and the questionnaire is not stable in its two-factor structure. Different factors were found with the EFA for the regions. The questionnaire appears to be culturally-sensitive.

### Strengths and limitations

A strength of this study is that statistical analyses were carried out on a larger sample than has been used in Spain. The questionnaire has been tested in three different Dutch regions, and hence was applied in different organisational circumstances within the healthcare system. The analyses were performed by region and can be perceived as three separate studies. The sample size in one of the regions was smaller than recommended.<sup>30</sup> In the other regions, a sample size was obtained that was close to the recommended sample size. Finally, the analysis of the remaining region and the analysis on the entire data set were well above the recommended sample size. Another strength of this study is the measurement of the test-retest reliability. To test whether measurement results are reproducible in test-retest situations is important.<sup>28,31</sup> No test-retest reliability analysis had been conducted for the original questionnaire. A possible explanation of the poor test-retest results could be current collaborative experiences influencing the opinions of the responding medical professionals.

A limitation of this study is that the response rate (33%), although higher than in the original study, was limited.

### Comparison with existing literature

The strong theoretical base of the questionnaire, which was derived from several qualitative studies by D'Amour *et al*, supports the importance of the 10 dimensions related to interprofessional collaboration.<sup>5,24,27</sup> But to measure the relational and organisational items with a questionnaire comprising 10 questions seems difficult. One study which measured how GPs and SCCs rate their collaboration did not take into account the organisation they worked in.<sup>21</sup>

### Implications for research and practice

Even though the surveyed healthcare professionals considered the questions based on these dimensions relevant for their work in the Dutch healthcare context, the questionnaire proved not to be valid in the Netherlands. The construct validity and test-retest reliability of this translated questionnaire were insufficient to warrant use in the Dutch setting. To develop a useful questionnaire about collaboration, more questions about each of the indicators are probably needed, which means constructing a new questionnaire and testing it with an approach similar to the one presented here. Alternatively, use of qualitative methods such as interviews could lead to more insight into interorganisational collaboration. Even though a pre-testing process might indicate a high face-validity, it is not a guarantee as to other aspects of validity. Even between regions in the same country, cultural differences can exist.

Importantly, this study demonstrates that even when the face-validity of a translated questionnaire seems to be good, thanks to a pre-testing process with sufficient opportunities for cultural adaptation, other aspects of validity must be tested before applying the instrument in new settings. Even between regions, there can be cultural differences. The authors advise caution when using a questionnaire in another cultural setting, population, or context without testing it. Before choosing a questionnaire for a small-scale study in general practice, the authors recommend checking that it is validated for that setting.

### Funding

The authors thank the Hein Hogerzeil Foundation, which has contributed to funding for this research.

### Ethical approval

Approval for this study was obtained by the Medical Ethics Review Committee (MERC) University Medical Center Utrecht.

### Provenance

Freely submitted; externally peer reviewed.

### Acknowledgements

The authors would like to thank Robert Nūno-Solis for providing detailed information about their study and for providing the authors with their dataset, in order to compare the different analyses.

## References

1. Barnett K, Mercer SW, Norbury M, et al. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. *Lancet* 2012; **380(9836)**: 37–43. doi: 10.1016/S0140-6736(12)60240-2
2. Cassel CK, Reuben DB. Specialization, subspecialization, and subspecialization in internal medicine. *N Engl J Med* 2011; **364(12)**: 1169–1173. doi: 10.1056/NEJMs1012647
3. Institute of Medicine (US) Committee on Quality of Health Care in America. *A New Health System for the 21st Century: Crossing the Quality Chasm*. Washington, DC: National Academies Press, 2001.
4. Starfield B. Primary care, specialist care, and chronic care: can they interlock? *Chest* 2010; **137(1)**: 8–10. doi: 10.1378/chest.09-1441
5. D'Amour D, Ferrada-Videla M, San Martin Rodriguez L, et al. The conceptual basis for interprofessional collaboration: core concepts and theoretical frameworks. *J Interprof Care* 2005; **19(Suppl 1)**: 116–131. doi: 10.1080/13561820500082529
6. Göbel B, Zwart D, Hesselink G, et al. Stakeholder perspectives on handovers between hospital staff and general practitioners: an evaluation through the microsystems lens. *BMJ Qual Saf* 2012; **21(Suppl 1)**: i106–i113. doi: 10.1136/bmjqs-2012-001192
7. Smith SM, Allwright S, O'Dowd T. Effectiveness of shared care across the interface between primary and specialty care in chronic disease management. *Cochrane Database Syst Rev* 2007;(3): CD004910. doi: 10.1002/14651858.CD004910.pub2
8. Akkerman SF, Bakker A. Boundary crossing and boundary objects. *Review of Educational Research* 2011; **81(2)**: 132–169. doi: 10.3102/0034654311404435
9. Hesselink G, Vernooij-Dassen M, Pijnenborg L, et al. Organizational culture: an important context for addressing and improving hospital to community patient discharge. *Med Care* 2013; **51(1)**: 90–98. doi: 10.1097/MLR.0b013e31827632ec
10. Yemm R, Bhattacharya D, Wright D, et al. What constitutes a high quality discharge summary? A comparison between the views of secondary and primary care doctors. *Int J Med Educ* 2014; **5**: 125–131. doi: 10.5116/ijme.538b.3c2e
11. Johnson JK, Farnan JM, Barach P, et al. Searching for the missing pieces between the hospital and primary care: mapping the patient process during care transitions. *BMJ Qual Saf* 2012; **21(Suppl 1)**: i97–i105. doi: 10.1136/bmjqs-2012-001215
12. Kripalani S, Jackson AT, Schnipper JL, et al. Promoting effective transitions of care at hospital discharge: a review of key issues for hospitalists. *J Hosp Med* 2007; **2(5)**: 314–323. doi: 10.1002/jhm.228
13. O'Malley AS, Reschovsky JD. Referral and consultation communication between primary care and specialist physicians: finding common ground. *Arch Intern Med* 2011; **171(1)**: 56–65. doi: 10.1001/archinternmed.2010.480
14. Berendsen AJ, de Jong GM, Meyboom-de Jong B, et al. Transition of care: experiences and preferences of patients across the primary/secondary interface — a qualitative study. *BMC Health Serv Res* 2009; **9**: 62. doi: 10.1186/1472-6963-9-62
15. Hassink-Franke LJ, Janssen MM, Oehlen G, et al. GPs' experiences with enhanced collaboration between psychiatry and general practice for children with ADHD. *Eur J Gen Pract* 2016; **22(3)**: 196–202. doi: 10.1080/13814788.2016.1177506
16. Chiocchio F, Lebel P, Dubé JN. Informational role self-efficacy: a validation in interprofessional collaboration contexts involving healthcare service and project teams. *BMC Health Serv Res* 2016; **16**: 153–153. doi: 10.1186/s12913-016-1382-x
17. Goldman J, Reeves S, Wu R, et al. A sociological exploration of the tensions related to interprofessional collaboration in acute-care discharge planning. *J Interprof Care* 2016; **30(2)**: 217–225. doi: 10.3109/13561820.2015.1072803
18. Bosch M, Dijkstra R, Wensing M, et al. Organizational culture, team climate and diabetes care in small office-based practices. *BMC Health Serv Res* 2008; **8**: 180–180. doi: 10.1186/1472-6963-8-180



19. Orchard CA, King GA, Khalili H, et al. Assessment of Interprofessional Team Collaboration Scale (AITCS): development and testing of the instrument. *J Contin Educ Health Prof* 2012; **32(1)**: 58–67. doi: 10.1002/chp.21123
20. Ohman-Strickland PA, John Orzano A, Nutting PA, et al. Measuring organizational attributes of primary care practices: development of a new instrument. *Health Serv Res* 2007; **42(3 Pt 1)**: 1257–1273. doi: 10.1111/j.1475-6773.2006.00644.x
21. Berendsen AJ, Benneker WH, Groenier KH, et al. DOC questionnaire: measuring how GPs and medical specialists rate collaboration. *Int J Health Care Qual Assur* 2010; **23(5)**: 516–526. doi: 10.1108/09526861011050547
22. Nuño-Solinis R, Berraondo Zabalegui I, Sauto Arce R, et al. Development of a questionnaire to assess interprofessional collaboration between two different care levels. *Int J Integr Care* 2013; **13**:e015. doi: 10.5334/ijic.984
23. D'Amour D, Goulet L, Labadie JF, et al. A model and typology of collaboration between professionals in healthcare organizations. *BMC Health Serv Res* 2008; **8(1)**: 188–188. doi: 10.1186/1472-6963-8-188
24. D'Amour D, Oandasan I. Interprofessionality as the field of interprofessional practice and interprofessional education: an emerging concept. *J Interprof Care* 2005; **19(Suppl 1)**: 8–20. doi: 10.1080/13561820500081604
25. Oandasan I, D'Amour D, Zwarenstein M, et al. *Interdisciplinary education for collaborative, patient-centred practice research and findings report*. Ottawa, ON: Health Canada. 2004.
26. World Health Organization. Process of translation and adaptation of instruments. [http://www.who.int/substance\\_abuse/research\\_tools/translation/en/](http://www.who.int/substance_abuse/research_tools/translation/en/) (accessed 30 Jan 2018).
27. Olsson U. Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika* 1979; **44(4)**: 443–460. doi: 10.1007/BF02296207
28. Fraley C. How many clusters? Which clustering method? Answers via model-based cluster analysis. *The Computer Journal* 1998; **41(8)**: 578–588. doi: 10.1093/comjnl/41.8.578
29. Fleiss JL, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educ Psychol Meas* 1973; **33(3)**: 613–619. doi: 10.1177/001316447303300309
30. Kline RB. *Principles and practice of structural equation modelling*. 2nd ed. New York, NY: Guilford, 2005.
31. Beatty PC, Willis GB. Research synthesis: The practice of cognitive interviewing. *Public Opin Q* 2007; **71(2)**: 287–311. doi: 10.1093/poq/nfm006