Health Research Policy and Systems

https://doi.org/10.1186/s12961-026-01507-z

Article in Press

Scoping review of instruments measuring organizational context in implementation studies within healthcare

Received: 30 Jan 2026

Accepted: 17 Jun 2026

Published online: 28 July 2026

Cite this article as: Squires, J., Graham, I., Santos, W. *et al.* Scoping review of instruments measuring organizational context in implementation studies within healthcare. *Health Res Policy Sys* (2026). https://doi.org/10.1186/s12961-026-01507-z

Janet Squires, Ian Graham, Wilmer Santos, Ligyana Candido, Vanessa Watkins, Megan Greenough, Alison Hutchinson, Chantal Backman, Anna Bergström, Melissa Brouwers, Christopher Burton, Christine Cassidy, Cheyne Chalmers, Heather Colquhoun, Janet Curran, Melissa Varin, Paula Doering, Annette Rose, Christina Godfrey, Doris Grinspun,

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

ARTICLE IN PRESS

©The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/

ARTICLE IN PRESS

Scoping Review of Instruments Measuring Organizational Context in Implementation Studies within Healthcare

Janet E. Squires[*,1,2], Ian D. Graham[1,2,3,4], Wilmer J. Santos[1,2], Ligyana Korki de Candido[1], Vanessa J. Watkins[4], Megan Greenough[1,5], Alison M. Hutchinson[4,6] and the ICON Team[7]

* Corresponding Author

1 University of Ottawa, School of Nursing, Ottawa, Canada

2 Ottawa Hospital Research Institute, Methods and Implementation Research Program, Ottawa, Canada

3 University of Ottawa, School of Epidemiology and Public Health, Ottawa, Canada

4 Deakin University, School of Nursing and Midwifery, Centre for Quality and Patient Safety Research in the Institute for Health Transformation, Geelong, Australia

5 Children's Hospital of Eastern Ontario, Ottawa, Canada

6 Barwon Health, Geelong, Australia

7 The ICON Team

ARTICLE IN PRESS

ARTICLE IN PRESS

Authors

Janet E. Squires, RN, PhD
Methods and Implementation Research Program, Ottawa Hospital Research Institute, Ottawa, Ontario, Canada
School of Nursing, Faculty of Health Sciences, University of Ottawa, Ottawa, Ontario, Canada
Email: janet.squires@uottawa.ca

Ian D. Graham, MA, PhD, FCAHS, NYAM, FRSC
Methods and Implementation Research Program, Ottawa Hospital Research Institute, Ottawa, Ontario, Canada School of Epidemiology and Public Health, School of Nursing, University of Ottawa, Ottawa, Ontario, Canada, Deakin University School of Nursing and Midwifery, Melbourne, Victoria, Australia
Email: igraham@ohri.ca

Wilmer J. Santos, RN, BScN, MScN
School of Nursing, Faculty of Health Sciences, University of Ottawa, Ottawa, Ontario, Canada
Methods and Implementation Research Program, Ottawa Hospital Research Institute, Ottawa, Ontario, Canada
Email: wsant020@uottawa.ca

Ligyana Korki de Candido, RN, MScN
School of Nursing, Faculty of Health Sciences, University of Ottawa, Ottawa, Ontario, Canada
Email: lkork013@uottawa.ca

Vanessa J. Watkins, RN, RM, PhD
Deakin University, School of Nursing and Midwifery, Geelong, Australia
Email: vanessa.watkins@deakin.edu.au

Megan Greenough, RN(EC), MScN, PhD

ARTICLE IN PRESS

ARTICLE IN PRESS

Children's Hospital of Eastern Ontario, Ottawa, Canada
University of Ottawa, School of Nursing, Ottawa, Canada
Email: mgreenough@cheo.on.ca

Alison M. Hutchinson, RN, BAppSci (Adv Nsg), MBioeth, PhD
Deakin University, School of Nursing and Midwifery; Centre for Quality & Patient Safety Research, Institute for Health Transformation, Geelong, Australia
Barwon Health, Geelong, Australia
Email: alison.hutchinson@deakin.edu.au

ICON team
Chantal Backman (Consortium representative)
University of Ottawa, School of Nursing, Ottawa, Canada
Email: Chantal.Backman@uottawa.ca

ARTICLE IN PRESS

ARTICLE IN PRESS

Abstract

Background

Healthcare providers' use of research in practice is suboptimal. In Learning Health Systems, using the best available research evidence to guide practice is fundamental to continuous improvement. Whilst there is a growing body of evidence that context modifies the effects of implementation strategies to promote research use, knowledge on how best to measure context is lacking. As part of a larger research program, which involved development of the Implementation in Context (ICON) framework comprising six domains and 22 attributes (Figure 1), we conducted a scoping review to identify psychometrically sound and pragmatic instruments for measuring macro- or meso-level ICON attributes.

Methods

The review was conducted in two phases: 1) identification of instruments that align with the 20 meso- and macro-level attributes in ICON, and 2) assessment of their psychometric and pragmatic

ARTICLE IN PRESS

ARTICLE IN PRESS

properties. Instruments and psychometric papers were identified through five electronic databases. The COnsensus-based Standards for selecting health Measurement INstruments (COSMIN) criteria guided the psychometric evaluation. A modified GRADE approach was used to summarize the quality of the evidence for each instrument. In this paper, we report the properties of the included instruments measuring one of the macro- or meso-level attributes.

ARTICLE IN PRESS

Results

We identified 38 instruments (and assessed 169 psychometric papers) that captured 9 of 15 meso-level ICON attributes; instruments measuring macro-level ICON attributes were not located. Most of the 38 instruments were rated as inadequate or doubtful (n=31) on content validity. Using the COSMIN recommendation ratings, 13 instruments are recommended for measuring 5 meso-level ICON attributes: organizational climate, organizational change processes, receptivity to change, evaluation activity, and leadership. The remaining 27

ARTICLE IN PRESS

instruments have the potential for use. All 38 instruments were rated to be at least adequate on pragmatic criteria.

Conclusions

We identified 13 instruments that measure meso-level ICON attributes that influence research implementation and meet psychometric quality criteria. Future work will include: 1) additional psychometric assessments of the 27 instruments not yet recommended for use, 2) psychometric and pragmatic assessments of instruments that assess multiple attributes in ICON, and 3) the development of instruments to measure ICON attributes for which instruments were not identified.

ARTICLE IN PRESS

Registration: Open Science Framework, 21 June 2022, https://doi.org/10.17605/OSF.IO/B2Z8K

Keywords - Delivery of Health Care, Implementation Science, Knowledge Translation, Implementation in Context (ICON) Framework, Measurement, Instrument, Psychometrics, scoping review, COSMIN

ARTICLE IN PRESS

Introduction

Healthcare providers' use of research in clinical practice is suboptimal, despite increased awareness of evidence and known gaps (1-9). In a recent knowledge synthesis of inappropriate healthcare across Canada, on average, 30% of clinical practices reported in 174 studies were *not* aligned with the best available evidence (9). Suboptimal levels of research-informed care have also been reported globally: USA (45%) (5); United Kingdom, Australia, and New Zealand (51%) (8); and Australia (43%) (7). In response, there is widespread interest in the science and practice of research implementation to advance Learning Health Systems, which aim to integrate clinical practice with research, enabling continuous learning and improvement in healthcare (10). Researchers have identified several major methodological issues facing the implementation field (11-16), including measuring and reporting the influence of context in the design, assessment, and outcomes of implementation strategies. Contextual factors can impact the

ARTICLE IN PRESS

ARTICLE IN PRESS

effectiveness of implementation strategies by exacerbating the problem the strategy is designed to overcome, influencing the susceptibility of the problem to the strategy, and determining how the strategy can work (17). Context differs between settings and changes over time, making its assessment critical to the design and delivery of optimally effective implementation strategies, facilitating evaluation of the transferability of implementation strategies between settings, and optimising the sustainability of implementation strategies (17). Whilst there is growing recognition that context can modify the effects of implementation strategies, we lack knowledge on how best to measure context (18, 19). Evaluating context within Learning Health Systems is crucial to identifying key questions and facilitating actions for continuous improvement. Authors of a recent realist synthesis found that effective evaluations depend on fostering supportive organizational conditions, rapid feedback loops, and a culture of innovation (20). Understanding these contextual factors informs research implementation and drives the sustainability of health innovation. To provide conceptual clarity regarding the attributes of context, we conducted a large body of

ARTICLE IN PRESS

research, which led to the development of the Implementation in CONtext (ICON) framework, described in the next section. To understand how contextual attributes operate in a specific setting within a Learning Health System, implementers and researchers require valid, reliable, and feasible measures. Use of such measures, importantly, can provide: an understanding of contextual attributes prior to implementation; inform strategies to address attributes identified as potential influences on implementation success; and an understanding of contextual attributes following implementation.

The Implementation in CONtext (ICON) Framework

Figure 1 depicts the ICON framework, which guided the project reported in this paper. Development of ICON was prompted by a lack of conceptual clarity about context in implementation. Thus, our earlier work focussed on the first steps to gain clarity about implementation context through a series of four studies. Study 1 was a concept analysis of context in implementation (N=70 articles) (21, 22). Study 2 was an analysis of 145 interviews with a diverse sample of healthcare providers, exploring elements of context relevant to their use of research findings

ARTICLE IN PRESS

in clinical practice (23). Study 3 comprised interviews with 39 health system leadership stakeholders in four high-income countries to elicit their perceptions of important elements of context in implementation efforts (24). Study 4 involved the triangulation of results from studies 1-3 to produce the ICON framework, comprising 3 levels, 6 domains, 22 core attributes, and 108 features of context relevant to implementation (25) (**Figure 1**). Definitions of the 22 core attributes of ICON are located in the manuscript detailing its development (25). More recently, we conducted a study that indicates all ICON attributes also apply in low- and middle-income countries (26). The ICON framework provides conceptual clarity on context in implementation, which we define as: the characteristics of the provider group and the healthcare recipient group, the internal organizational characteristics (climate and structures, social behavior, response to change, and processes), and the external influences of the broader healthcare system. These contextual domains have the potential to directly or indirectly influence or modify the use of research evidence by healthcare providers, policymakers and recipients of care. ICON is the only implementation meta-framework solely dedicated to

ARTICLE IN PRESS

the role of context in implementation. While context is included in several other implementation frameworks (e.g., Promoting Action on Research Implementation in Health Services (PARIHS) Framework (27, 28), Knowledge-to-Action Framework (KTA) (29), the Consolidated Framework for Implementation Research (CFIR) (30, 31), it is not sufficiently detailed in these frameworks. For instance, many of the features of context attributes reported in ICON are not identified in other implementation frameworks; examples include technology (attribute *Resource Access),* mentorship (attribute *Leadership*), continuity of care (attribute *Work Structure),* funding model (attribute *Financial),* social interactions (attribute *Collaboration*), accountability (attribute *Professional Role),* and system complexity (attribute *System Features)* (32).

ARTICLE IN PRESS

Previous Measurement Reviews on Context

A small number of syntheses (33-41) conducted by other research teams have identified context instruments, but they are based on either incomplete conceptualizations of context and/or were limited to a single setting. For example, Chaudoir and colleagues(33) assessed measures

ARTICLE IN PRESS

for six ICON attributes (organizational culture, physical and technological resources, economic arrangements, governance, leadership, organizational climate), limited to quality improvement studies. Brennan et al.(34) assessed measures for seven ICON attributes (service provider population, organizational climate, organizational culture, leadership, physical and technological resources, communication processes, receptivity to change), in primary care only. Rogers et al.(35) reported a synthesis of 64 studies on context measurement; most instruments measured a single ICON attribute or were developed for specific innovations and are, therefore, not suitable for generic context assessments in implementation across innovations or healthcare sectors. More recently, the Society for Implementation Research Collaboration (SIRC) assessed measures that map onto the CFIR framework, which categorizes and describes contextual determinants across five domains that may influence implementation (30, 31). Their work is published in five reviews, each targeting selected constructs in four of the five context-related CFIR domains: (1) outer setting (e.g. cosmopolitanism, peer pressure) (37), (2) inner setting (e.g.,

ARTICLE IN PRESS

leadership engagement, organizational culture) (38-40), (3) characteristics of individuals (e.g. knowledge and beliefs about the innovation, self-efficacy) (41), and (4) implementation processes (e.g. planning, engaging) (39). Although these reviews significantly expanded the knowledge base on how to measure context, several gaps remained. First, while CFIR includes 18 constructs related to context, all but one (culture) are features (examples) of the 22 ICON context attributes. Altogether, the 18 constructs in CFIR represent only 28% (38 of 136) of the context features/constructs that comprise the 22 context attributes in ICON. Therefore, many additional instruments not identified in the SIRC reviews will exist. Second, the SIRC reviews are limited to a single healthcare setting (mental health). Third, the mapping of instruments to context constructs was based on each instrument's title/label. This can result in inaccurate mapping. For example, the Patient Safety Climate instrument (42) was mapped to climate; however, the instrument items measure cultural norms, and thus, this instrument measures culture, not climate. A more appropriate approach is to map instruments based on a review of their included items. Finally, in the

ARTICLE IN PRESS

ARTICLE IN PRESS

SIRC project the reviewers assessed psychometric properties using the Psychometric and Pragmatic Evaluation (PAPERS) tool (43). While this tool assesses nine psychometric criteria, it omits the most important validity criterion of content validity: the degree to which instrument items reflect the construct(s) to be measured (44). Content validity is fundamental to assessing the validity of an instrument because lack of it can decrease internal consistency, structural validity, and interpretability of scores obtained from the instrument (45).

In this paper, we report a scoping review of instruments measuring meso- and macro-level attributes of context important to implementation by healthcare professionals, as captured in the ICON framework. We undertook a scoping review to map the breadth and depth of relevant instruments and to highlight gaps in the available instruments(46). Our purpose was to identify psychometrically sound and pragmatic instruments for measuring macro- or meso-level ICON attributes. The review objectives were to: 1) systematically search and critically review published and unpublished literature for quantitative instruments of meso-level (organizational) and macro-level (external to the

ARTICLE IN PRESS

ARTICLE IN PRESS

organization, including between organizations) context attributes used in healthcare implementation efforts, and 2) assess the development, psychometric (reliability, validity), and pragmatic (acceptability, easiness) properties of the identified instruments across healthcare sectors. In this paper, we report the properties of the included instruments measuring one of the macro- or meso-level ICON attributes.

Methods

Design

In the reporting of this scoping review, we adhered to the PRISMA Extension for Scoping Reviews (PRISMA-ScR) (46) and the PRISMA-COSMIN for Outcome Measurement Instruments guideline - a consensus-based reporting standard for reviews of health measurement tools (47)(**Additional File 1)**. The protocol is available in the Open Science Framework(48). We conducted the review in two stages: 1) identification of instruments that align with ICON attributes and mapping them to the 20 meso- and macro-level attributes in ICON, and 2) assessment of the psychometric and pragmatic properties of the identified instruments.

ARTICLE IN PRESS

ARTICLE IN PRESS

Stage 1- Identification of Instruments and Mapping to the ICON's Context Attributes

An experienced information specialist (TR) developed and applied a search strategy to five electronic databases (OVID MEDLINE, OVID PsycInfo, Ebsco CINAHL, Ovid Health and Psychosocial Instruments, and Wiley Cochrane Library) on October 6, 2020, which was PRESS reviewed(49). A filter for reviews (50, 51) was applied due to the large number of citations captured by the search strategy (N=17,876). Grey literature was not specifically searched; however, dissertations retrieved during the searches were reviewed. The search strategy and peer-review assessment are in **Additional File 2.**

ARTICLE IN PRESS

We used EndNote (52) to deduplicate records before uploading into DistillerSR (53) for screening and data extraction. We included studies that used an instrument to measure one or more ICON attribute(s). We excluded qualitative studies, study protocols, conference abstracts, editorials, opinion papers, case studies, and studies not published in English (**Table 1**). Two team members (DS, LA, MDV or WJS) independently screened the title and abstract of each record. Test

ARTICLE IN PRESS

screening of 10% of the full-text articles (n=100 articles) was conducted by two reviewers (DS, JZ, LA, MDV, or WJS) until a concordance rate of greater than 80% was achieved. The remaining records were equally divided among the reviewers for full-text screening.

Data extraction and full-text screening occurred simultaneously. Data extraction was performed by one reviewer (DS, JZ, LA, MDV, or WJS) and verified by a second reviewer (DS or LA). We evaluated all quantitative primary studies in each included review for an instrument that potentially measured any ICON context attribute. We used a standardized and pre-tested data extraction form to extract the instrument name, the cited development article for the instrument and the source of the instrument. For development articles in which instrument items were not reported, Google searches were performed to locate reports of the items. We excluded instruments that were not publicly available.

Two or more team members (AMH, IDG, JES, JZ, LA, LKC, MDV, and WJS) mapped

ARTICLE IN PRESS

each individual item of an included instrument to an ICON framework: 1) domain, 2) attribute, and 3) feature (if possible). Conflicts were resolved through consensus or consultation with senior team members (AMH, IDG, and JES). Instruments were divided as follows: 1) instruments composed of items measuring a single ICON attribute (e.g., leadership); 2) instruments composed of subscales measuring different ICON attributes; and 3) instruments composed of individual items measuring different ICON attributes. Because of the very large number of instruments identified, in this review, we report only on the first category - instruments composed of items measuring a single ICON attribute.

ARTICLE IN PRESS

Stage 2 – Search and Assessment of Psychometric and Pragmatic Properties

Search for Psychometric Articles of the Identified Instruments

From inception until August 2022, instruments were searched in PubMed using the Outcomes Measurement Properties (OMP) search filter (54). OMP has been validated as a precise and sensitive method for searching for psychometric articles on Patient Reported Outcomes

ARTICLE IN PRESS

Measures (PROMs) (54) and non-PROM instruments in PubMed (55). If the OMP search filter did not yield relevant articles and/or the instrument did not have a name, we performed a citation search in Scopus and Web of Science using the instrument's development article retrieved during Stage 1 and simultaneously conducted a forward search.

Eligibility Criteria, Article Selection, Data Extraction

We included 1) articles that reported on the validity and reliability of the included instruments used in any healthcare setting and sector; and 2) articles describing the development of instruments, regardless of the context (i.e., even if they did not originate in healthcare). We excluded synthesis studies, study protocols, conference abstracts, editorials, opinion papers, case studies, articles not published in English, articles not available in full-text, or studies that did not evaluate the instrument's psychometric properties (**Table 1**). Additionally, articles reporting instruments for which permission and payment was required were excluded.

We used Endnote (52) to deduplicate records before uploading them to DistillerSR (53) for screening and data extraction. Two team

ARTICLE IN PRESS

members (LKC, MG, VJW, or WJS) independently completed title and abstract screening, followed by full-text screening of each record. We developed a standardized data extraction form according to the COSMIN (COnsensus-based Standards for the selection of health Measurement INstruments) checklist (56, 57), which is considered best practice in the conduct of systematic reviews. Our data extraction form included: 1) instrument description, 2) study characteristics; 3) study setting; 4) sample demographics; 5) psychometric property evidence and rating (sufficient, insufficient, indeterminate, and inconsistent) for: a) content validity; b) structural validity; c) internal consistency; d) cross-cultural validity and measurement invariance; e) reliability; f) measurement error; g) hypothesis testing (discriminant, convergent, discriminative, and known groups validity); and h) responsiveness. Data extraction was performed independently by two team members (LKC, MG, VJW, or WJS) and in duplicate for 10% of included articles to ensure comprehensiveness of the data extraction form. After the first 10%, one reviewer extracted data from the remaining articles (LKC, MG, VJW, or WJS), which was cross-checked by a second reviewer

ARTICLE IN PRESS

ARTICLE IN PRESS

(WJS or MG). Conflicts in screening and data extraction were resolved through consensus or consultation with a senior researcher (AMH, IDG, JES).

Assessment of Measurement (Psychometric) Properties

In accordance with the COSMIN guidelines, a descriptive synthesis of the extracted measurement properties was performed, comprising psychometric properties and quality of evidence (56, 57). For each instrument, we used the COSMIN criteria to evaluate development and content validity, starting with the development article and then all subsequently published articles that evaluated all criteria or any single content validity criterion (relevance, comprehensiveness, and comprehensibility) (56). We ascribed a rating of sufficient (+), insufficient (–), inconsistent (±), or indeterminate (?) for each content validity criterion, guided by the COSMIN manual (56, 57). We followed this process for all other psychometric properties included in COSMIN and reported when an article utilized a variant or modified version of an instrument. This evaluation was conducted by one team member (LKC, MG, VJW, or WJS) and verified by a second reviewer (LKC, MG, VJW,

ARTICLE IN PRESS

ARTICLE IN PRESS

or WJS). Conflicts were resolved through consensus or consultation with a senior researcher (AMH, IDG, JES).

Synthesis of the psychometric findings was organized by healthcare sectors in accordance with COSMIN recommendations, with a threshold of 85% (56) for validity evaluations and 75% for the remaining psychometric properties (57). We then pooled psychometric properties based on healthcare sectors, regardless of the instruments' versions (i.e., some instruments had different versions).

To assess the quality of the psychometric evidence reported, we first completed the risk of bias questions outlined in the COSMIN manual for each psychometric property (from inadequate to very good) (56, 57). We then followed the GRADE approach, modified for COSMIN(57), to summarize the quality of evidence for the psychometric properties of each instrument in a particular healthcare sector (graded from very low to high). Based on the modified GRADE approach (57), the quality of evidence was initially assumed to be high, and then lowered based on evidence of risk of bias, inconsistency, imprecision, and indirectness.

ARTICLE IN PRESS

ARTICLE IN PRESS

The final component of our psychometric synthesis was to categorize the recommendation level of each instrument as A, B, or C as prescribed by COSMIN. Instruments in Category A have evidence of sufficient content validity and internal consistency and are recommended for use. Category B instruments require further psychometric testing but may be recommended for use. Category C instruments have strong evidence of insufficient measurement properties and should not be recommended for use (57).

Assessment of Pragmatic Properties

The pragmatic properties of the included instruments were evaluated and synthesized according to criteria outlined by Stanick and colleagues (43) on a -1 to 4 rating scale (with -1 being poor and 4 being excellent). The criteria used addressed acceptability (including the cost to use the instruments and use of accessible language) and easiness (including assessor burden and the length of the instrument). One team member (LKC, VJW, or WJS) evaluated the objective pragmatic criteria for the included instruments, and a second team member (LKC, VJW, or

ARTICLE IN PRESS

WJS) verified them. Conflicts were resolved through discussion or consultation with a senior researcher (AMH, IDG, JES).

Results

Stage 1- Identification of Instruments and Mapping to the ICON's Context Attributes

Our Stage 1 search using the review search filter resulted in 1132 records, of which 499 were selected for full-text screening (**Figure 2**). We included 194 reviews in our final Stage 1 sample, from which we identified 748 potential instruments. Two hundred and fifty-one of these instruments could be used without permission and at no cost; of these, 180 were mapped to an ICON attribute. In total, 38 instruments were homogenous with respect to ICON attributes, meaning they were composed of items that mapped to a single ICON attribute. All 38 instruments mapped to a meso-level ICON attributes; no instruments mapped to macro-level ICON attributes. **Figure 2** is a PRISMA diagram summarizing instrument identification and ICON mapping.

The 38 homogenous instruments included in this review captured 9 of the 15 meso-level ICON attributes, including: 1) organizational

ARTICLE IN PRESS

climate (n=3, 7.9%); 2) physical and technological resources (n=3, 7.9%); 3) internal relationships (n=6, 15.8%); 4) organizational culture (n=3, 7.9%); 5) organizational change processes (n=1, 2.6%); 6) receptivity to change (n=2, 5.1%); 7) evaluation activity (n=11, 28.9%); 8) leadership (n=7, 18.4%), and; 9) organization of work (n=2, 5.3%). **Figure 3** illustrates the distribution of instruments amongst the macro- and meso-level ICON framework attributes. In **Table 2,** we present the 38 instruments according to the four ICON framework meso-level domains and the 9 attributes targeted in the instruments, including descriptions of instrument origins, the original and translated languages of the instruments, and the number of subscales and items for each instrument, based on their development articles. The table is organized first by ICON domain and then by each domain's respective attributes.

ARTICLE IN PRESS

Stage 2 – Search and Assessment of Psychometric and Pragmatic Properties

Search Results

Our Stage 2 searches yielded 17,450 records reporting psychometric properties across the 38 included instruments. We

ARTICLE IN PRESS

screened 16,377 records after removing duplicates and assessed 623 full-text articles for inclusion (**Figure 4**). Included were 169 articles, with more than half (n=102, 60.4%) published between 2012 and 2022 (**Additional File 3** lists excluded articles and the reasons). Most articles were published in Europe (n=68, 40.2%), followed by North America (n = 65, 38.5%), primarily in the United States of America (n=51, 30.2%), then Asia (n=28, 16.6%), South America (n=8, 4.7%), Oceania (n=7, 4.1 %), and Africa (n=1, 0.6%).

Assessment of Instrument Development

In **Table 3**, we present the findings on the quality of development of the 38 instruments using each instrument's development study. We located the development study for most (n=35, 92%) instruments, except for the following: 1) Questionnaire on Computer Systems and Decision Making(58), 2) American Board of Internal Medicine's (ABIM) Questionnaire(59, 60), and 3) Leadership Behavior Description Questionnaire(61). Although, we didn't locate the development studies for these three instruments, we provided references in Table 3 that describe them. In total we included 37 development

ARTICLE IN PRESS

ARTICLE IN PRESS

articles in our assessment of 35 instruments, as the Team Climate Inventory (TCI) (62, 63) and the Interpersonal processes of care: IPC - 29 (64, 65) had two articles each describing their development. Based on COSMIN guidelines(56), for content validity, most instruments were rated as either inadequate (n=16, 42.1%)(66-81) or doubtful (n=15, 39.5%)(42, 62-65, 73, 82-93). Of the instruments with a development study, two (5.3%) were rated to have very good evidence of content validity: 1) Community Impacts of Research Oriented Partnerships (CIROP)(94) and 2) CollaboRATE Measure(95); while two (5.3%) were rated to have adequate evidence of content validity: 1) Team Decision Making Questionnaire(96) and 2) Care Transitions Measure (CTM)(97).

ARTICLE IN PRESS

Articles reporting the development study for 35 of the included instruments clearly

described the intended construct, the target population for which the instrument was developed and the context for its use. For most instruments, the gaps in content validity were due to a lack of detailed reporting of the methods used and/or results for the items' relevance

ARTICLE IN PRESS

(n=30, 78.9%), comprehensibility (n=29, 76.3%), and comprehensiveness (n=29, 76.3%). Further, only half of the development studies (n=19, 50%) included a cognitive interview or pilot study with the target sample relevant to the instruments' origin. In 16 (42.1%) of the instruments, the sample was not described in sufficient detail and/or included only topic experts and not the target population for which the instrument was developed. See **Table 3** for further details of our content validity assessment of each instrument.

Measurement (Psychometric) Properties of the 38 Instruments

In **Table 4**, we summarize the psychometric properties of each instrument, organized first by ICON attribute and second by the healthcare sectors in which they were evaluated. In the table, we include the number of articles reporting a psychometric property, the overall rating of each psychometric property, and the quality of evidence pertaining to the psychometric rating. Next, we provide a descriptive summary of the data, organized according to the targeted ICON attribute. In total, we summarized 169 articles for the 35 instruments that we were able to locate psychometric studies for. There is variability in

ARTICLE IN PRESS

ARTICLE IN PRESS

the number of psychometric articles for each of the instrument (ranging from one article to 35 articles) and across the sectors in which instruments were tested in which are noted in **Table 4**.

Organizational Climate. Three instruments (62, 63, 66, 82) assess *Organizational Climate* (**Table 4**). Organizational climate is defined as " to shared recurring patterns of behaviour, attitudes and feelings which characterize life in that organization"(25). We mapped the three instruments into the organizational climate attribute because they measured aspects of team climate which is a feature of organizational climate within the ICON framework (25) (see **Table 2** for the feature related to each instrument). The following instruments are appropriate to use in the respective healthcare sectors: 1) Internal Participation Scale (66) (rehabilitation); 2) Team Climate Inventory (TCI) (62, 63, 98) (acute care/specialty care and primary care), and; 3) Costa and Anderson's (2011) (82) Measure of Trust in Teams (acute care/specialty care). While additional studies evaluated the psychometric properties of the TCI in home/community/public health settings (99, 100), content validity was not appraised in these sectors; therefore, the TCI cannot be

ARTICLE IN PRESS

ARTICLE IN PRESS

recommended for use in these sectors at this time. Further, several variants of the TCI were evaluated; 3 of 8 (37.5%) studies evaluated the short form in primary care (101), home or community care (102), and public health (103) sectors. Two of these studies confirmed, using confirmatory factor analysis, the same four-factor structure (participative safety, support for innovation, vision, and task orientation) composed of 14-items (102, 103). Beaulieu and colleagues found a similar factor structure but only conducted an exploratory factor analysis in testing a 19-item short version of the TCI (101). The remaining five articles reported evaluations of a long version (38 items) of the TCI in acute/specialty care (62, 63, 98), primary care (104), and public health (100) settings.

Physical and Technological Resources. Three instruments (58, 67, 68) assess *Physical and Technological Resources* (**Table 4**). Physical and Technological Resources is defined as physical structures and technological resources of an organization that are required to deliver services (25). The instruments we mapped into this attribute were included because they measured technology, equipment and supplies or

ARTICLE IN PRESS

measured online resources which are features of the physical and technological resources ICON attribute (25). No psychometric papers were found for the Computer Systems and Decision Making instrument (58); thus, it could not be assessed to make a recommendation. The remaining two instruments (67, 68, 105) have the potential to be recommended for use; however, confirmatory factor analysis was not conducted to evaluate their theoretical factor structure, as required by the COSMIN guidelines. Therefore, further research is required.

Internal Relationships. Six instruments (69-71, 83, 94, 96) assess *Internal Relationships* (**Table 4**). Internal relationships are defined as "the ways in which two or more
people or groups within an organization regard and behave toward each other"(25). We attributed the six instruments to measure internal relationships because they pertained to partnerships or collaborative practices which is a feature of internal relationship in the ICON framework (25) All instruments have the potential to be recommended for use following further research. However, the Shared Decision-Making Questionnaire-Doc (SDM-Q-Doc) (71) can be recommended for

ARTICLE IN PRESS

ARTICLE IN PRESS

use in long-term care, as there were sufficient (although low quality) evidence of content validity and sufficient (moderate quality) evidence of internal consistency reported in two studies in long-term care (71, 106).

Organizational Culture. Three instruments (42, 84, 85) assess *Organizational Culture* (**Table 4)**. Organizational culture is defined as "the normative beliefs and shared expectations that govern the work behaviour of an organization" (25). We mapped the three instruments to organizational culture because they measured organizational norms which is a feature of organizational culture in the ICON framework (25) All three instruments had evidence of sufficient content validity (low quality) reported in a single article but lacked sufficient evidence of internal consistency to be recommended. An exploratory factor analysis was performed for two of the instruments (42, 84), and no internal consistency evidence was reported for the final instrument (85). Therefore, all three instruments (42, 84, 85) have the potential to be recommended for use but first require more research.

ARTICLE IN PRESS

ARTICLE IN PRESS

Organizational Change Processes. One instrument, Measure of Champion Behaviour (107), assesses *Organizational Change Processes* (**Table 4)**. Organizational change processes is defined as the process of altering an organization's strategies, processes, procedures, technologies and/or culture to improve service delivery" (25). We mapped Measure of Champion Behaviour (107) into the organizational change processes because champions/opinion leaders is a feature of this ICON attribute (25). This instrument had sufficient evidence of high-quality content validity, structural validity, internal consistency, and convergent validity, and is recommended for use. However, it was only evaluated in one psychometric study conducted in business settings, not in healthcare.

Receptivity to Change. Two instruments (73, 86) assess *Receptivity to Change* (**Table 4**) and are suitable for use in selected sectors. Receptivity to change is defined as "an openness and responsiveness to ideas, impressions or suggestions, and a readiness or fit of critical features

of the environment as they specifically relate to a targeted practice" (25). We mapped the two instruments to the receptivity to change attribute

ARTICLE IN PRESS

because they measured compatibility or readiness for change which are features of this ICON attribute(25). The Organizational Readiness for Implementing Change (ORIC) Scale (86) is recommended for use based on its development article, which reported assessments in multiple settings and sectors from acute care to university students and non-government organizations. However, in other studies, there is uncertainty in the evidence related to content validity of the ORIC in acute care (n=2 studies (108, 109)) and primary care (n=1 study (110)). In studies in the home/community sector (111) and primary care (110, 112), researchers found poor data fit with the two-factor structure theorized and tested in the development study (86).

The Resistance to Change Scale (73) is recommended for use based on one study (73) showing sufficient content validity (low quality) evidence, and two studies (73, 113) identifying sufficient internal consistency and structural validity (high quality) evidence of a four-factor structure, but neither was conducted in healthcare. One of these studies demonstrated sufficient measurement invariance (high-quality) across data from 17 countries (113). The other study showed sufficient

ARTICLE IN PRESS

discriminative validity between resistance to change and measures related to personality traits and cognitive ability, and demonstrated the ability of the 'resistance to change' score to discriminate between individuals who are more likely to change behaviour after an imposed system change (73).

Evaluation Activity. Eleven instruments (59, 60, 64, 65, 74, 87-92, 95, 97) assess *Evaluation Activity* (**Table 4**), four of which are recommended for use: 1) Interpersonal Processes of Care (IPC) (64, 65); 2) Shared Decision-Making Questionnaire (SDM-Q-9) (90); 3) Patient Assessment of Chronic Illness Care Survey (PACIC) (91); and 4) Care Transitions Measure (CTM) (97). Evaluation activity is defined as "the systematic collection of information

about the activities, characteristics and outcomes of programs, services, policies or processes in an organization, to make judgements about the program/process, improve effectiveness and/or inform decisions about future development in that organization"(25). Although the instruments we mapped into the evaluation activity attribute may relate to other ICON attributes (e.g., measures of shared decision making could also

ARTICLE IN PRESS

ARTICLE IN PRESS

relate to internal relationships), the 11 instruments we included in the evaluation activity are intended to be used to garner patient/client/consumer feedback to staff regarding the services they received, which is a feature of the evaluation activity ICON attribute(25).

Different versions of the Interpersonal Processes of Care (IPC) instrument were developed based on the original conceptual framework and items (64). Variants of the IPC I version (41-57 items) were evaluated in two studies, one in public health (64) and the other in a home/community setting(114). Variants of the IPC II version (16-29 items) were evaluated in seven studies, three in acute/specialty care (65, 115, 116) and four in primary care (117-120). The IPC II can be recommended for use in acute/specialty care (65, 115, 116) and in primary care (117-120), but there is insufficient evidence to recommend the IPC I for use in the home/community (114) and public health settings (64).

In a single study (90), the original SDM-Q questionnaire with 11 items was evaluated in both primary and acute care settings and is

ARTICLE IN PRESS

recommended for use in both settings. The remaining 21 studies evaluated a variant of the most updated version of the instrument, the SDM-Q-9 (7-9 items), in either acute/specialty care (n=16 articles (121-135), primary care (n=4 articles (136-139), or both acute/specialty care and primary care (n=1 article (140)). The SDM-Q-9 is recommended for use in both acute/specialty care (121-135, 140) and primary care (121-135, 140).

The original PACIC (20 items) was evaluated in 24 studies (13-20 items) in acute or specialty care (n=11 articles (141-151)), primary care (n=11 articles (91, 152-161)), rehabilitation (n=1 article (162)), and in an unidentified sector (n=1 article (163)). The short version of the PACIC (11 items) was evaluated in 6 studies in acute/specialty care (n= 4 articles (143-145, 164)) primary care (n=2 articles (152, 165)). In one stud y(166), the short version of the PACIC was evaluated with a mix of individuals from acute/specialty care, primary care, long-term care, home/community, and public health sectors. An extended version of the PACIC (PACIC5A), with six additional items that evaluate actions providers can perform to facilitate behaviour change in patients, was

ARTICLE IN PRESS

ARTICLE IN PRESS

evaluated in acute/speciality care (n=3 articles (167-169)) and primary care (n=3 articles (170-172)). Finally, a study conducted in primary care (154) evaluated a 27-item version of the PACIC, which included seven additional items assessing culturally sensitive care. The original PACIC is recommended for use in acute/specialty care (141-151) and in primary care (91, 152-161), whereas the short version of the PACIC is recommended for use in acute/specialty care (143-145, 164) and in primary care (152, 165).

Seven studies (173-179) evaluated the psychometric properties of the CTM-3 in acute/specialty care, and 12 studies evaluated the psychometric properties of the CTM-15 in acute/specialty care (n=10 articles (173-182)), primary care (n=1(97)), and rehabilitation (n=1 (183)). The CTM-3 and CTM-15 are recommended for use in acute/specialty care. However, it is important to note that despite the informed COSMIN recommendation, there was variability in the psychometric properties reported across the articles. For instance, while four articles (174-176, 180) reported that the original four-factor

ARTICLE IN PRESS

structure sufficiently fit the data, four other articles (178, 181, 182) did not.

Leadership. Seven instruments (61, 75-80) assess *Leadership* (**Table 4)**, three of which are recommended for use: 1) the Implementation Leadership Scale in acute/specialty care (n=7 articles (77, 184-189)); 2) the Servant Leadership Questionnaire in a non-healthcare setting (n=1) (78); and 3) the Servant Leadership Survey also in non-healthcare settings (n=3 articles) (79, 190, 191). Leadership is defined as "refers to the types and styles of leaders within an organization"(25). We mapped six instruments to the leadership attribute because they measured leadership styles which is a feature of the leadership attribute in the ICON framework (25);t he Mentorship Effectiveness Scale (80) pertained to the mentor feature of the leadership ICON attribute (25). A lack of evidence regarding the content validity of the four remaining leadership instruments precludes recommending their use.

Organization of Work. Two instruments assess *Organization of Work* (**Table 4)**: 1) NASA Task Load Index (81) and 2) RAFAELA

ARTICLE IN PRESS

System, which consists of the Professional Assessment of Nursing Care Intensity Level (PAONCIL) and Oulu Patient Classification instrument (OPC) (93). Organization of work is defined as "refers to arrangement of tasks, responsibilities and resources within and between service providers working in the organization" (25). We mapped the two instruments to the organization of work attribute because they measured workload which is a feature of this ICON attribute (25). We determined that both instruments have the potential to be recommended for use with more psychometric testing. The 6-item unidimensional NASA Task Load Index was reported to have sufficient structural validity, internal consistency, and divergent and convergent validity when tested in diabetes management care(192) and in 18 intensive care units (193). With the RAFAELA System, a study (194) conducted in a single hospital demonstrated the relevance of the OPC as a measure of nursing intensity based on the care required by patients, and another article (93) described the importance of using both an objective (OPC) and subjective (PAONCIL) measure of nursing intensity. In the other studies, researchers aimed to evaluate the minimum requirements (e.g.,

ARTICLE IN PRESS

ARTICLE IN PRESS

explanatory power and response rate) required for the RAFAELA system to accurately measure nursing intensity (195), and hypothesis testing between two hospitals (196), medical and surgical units (197), and liver versus kidney transplant patients (198). Although the subscales of the OPC were reported to have $\alpha \geq 0.70$ in two studies (194, 199), no articles reported sufficient structural validity.

Pragmatic Properties of the 38 Instruments

Additional File 4 shows the detailed pragmatic assessment of the 38 instruments. Using the PAPERS' criteria (43), none of the instruments were rated as excellent for all pragmatic properties pertaining to acceptability and easiness. However, all the instruments (42, 58-97) were rated as at least adequate for all four areas assessed, including those measuring acceptability (cost and language accessibility) and easiness (assessor burden for training and interpretation and length). Regarding accessibility, all instruments (42, 58-97) were found to be cost-free and/or available in the public domain and were rated as excellent in this criterion. Multiple instruments were rated to have adequate accessibility (n=16, 42.1%) (42, 58, 62, 63, 66-69, 72, 75-77,

ARTICLE IN PRESS

ARTICLE IN PRESS

80, 84, 89, 93, 96), with readability at college-level. Thirteen instruments (34.2%) were rated to have excellent accessibility (61, 64, 65, 73, 74, 78, 79, 82, 86, 88, 91, 92, 95, 97), with a readability level below the 8th grade. Finally, nine instruments (23.7%) were rated as having good language accessibility (59, 60, 70, 71, 81, 83, 85, 87, 90, 94), with a readability level between the 8th and 12th grades. In relation to *easiness*, 35 (92.1%) instruments required no specific training for their administration, and were rated as excellent (42, 58-60, 64-92, 94-97). The TCI(62, 63), Leadership Behavior Description Questionnaire(61) and the RAFAELA system(93) were rated as good since the instruments included a free or minimal-cost manual for self-training. All instruments except the RAFAELA system(93) (n=37, 97.4%) (42, 58-92, 94-97) were rated to have adequate assessor burden in terms of interpretation. Lastly, most instruments (n=29, 76.3%) had between 10 and 50 items (42, 58-65, 68-70, 72-75, 77-80, 82-89, 91, 92, 94, 96, 97), which is considered a good length, and 9 instruments (23.7%) had excellent length with 10 or fewer items (66, 67, 71, 76, 81, 90, 93, 95).

ARTICLE IN PRESS

Discussion

Summary of Findings

ICON comprises three levels (macro, meso and micro), 6 domains, and 22 attributes (25). In this study, we reviewed instruments that measure a single ICON meso-level attribute and did not locate any instruments that mapped to ICON macro-level attributes. We identified 38 instruments that measured 9 of the 15 meso-level ICON framework attributes of context (25). The number of instruments per attribute ranged from 1 (for the attribute *Organizational Change Processes*) to 11 (for the *Evaluation* attribute). Most instruments were: 1) developed in the United States (n=21, 55.3%); 2) written originally in English (n=34, 89.5%); 3) tested initially in acute care or specialty care settings (n=18, 47.4%), primarily in acute care hospitals (n=15, 39.4%); and 4) targeted healthcare service providers (23 of 38, 60.5%). Further, we identified, assessed using the COSMIN guidelines, and reported on 169 studies with the primary aim of evaluating the psychometric properties of the 38 instruments. Most articles were published between 2012 and 2022.

ARTICLE IN PRESS

ARTICLE IN PRESS

Using COSMIN's psychometric recommendation ratings, 13 (34.2%) instruments measuring the following five ICON framework meso-level attributes met our criteria (could state them here) for recommending their use (Recommendation A – **Table 4**) in respective health sectors: For *Organizational Climate*: the 1) Internal Participation Scale in rehabilitation; 2) TCI in acute care/specialty care and primary care, and; 3) Measure of Trust in Teams in acute care/specialty care. For *Organizational Change Processes*: 1) Measure of Champion Behaviour in business settings. For *Receptivity to Change*: 1) ORIC scale in acute/specialty care and international non-government organizations, and 2) the Resistance to Change Scale in academic settings. For *Evaluation Activity*: 1) the CTM in acute/specialty care; 2) the IPC in acute/specialty care and primary care; 3) the SDM-Q-9 in acute/specialty care and primary care; 4) the PACIC in acute/specialty care and primary care. For *Leadership*: 1) the Implementation Leadership Scale in acute/specialty care; 2) the Servant Leadership Questionnaire in non-healthcare settings; and 3) the Servant Leadership Survey in non-healthcare settings. The remaining instruments with identified

ARTICLE IN PRESS

ARTICLE IN PRESS

psychometric articles (n=22, 57.9%) have the potential to be recommended for use but first require further psychometric testing (Recommendation B – **Table 4**). Barriers to recommending these instruments currently relate to: 1) limited evidence of content validity, including: i) the instrument has not been subjected to an assessment of content validity; ii) lack of description of the sample and inclusion of the potential users of instruments during cognitive testing or pilot testing; iii) lack of description of the methods used to evaluate relevance, comprehensiveness, and comprehensibility; and iv) lack of detailed findings related to content validity); and 2) lack of theory-informed evaluation of the instrument's structural validity (e.g., confirmatory factor analysis).

ARTICLE IN PRESS

In terms of pragmatic properties, all 38 instruments meet some criteria for acceptability and easiness (**Additional File 4**).

Content Validity

Borsboom et al. (200) and De Vet et al. (201) argue that content validity is the most important type of validity. De Vet et al (201) recommends, sequential evaluation of content validity, as instruments

ARTICLE IN PRESS

are applied within different contexts and/or for different purposes. Establishing content validity is primarily a qualitative process that involves: 1) determining a clear theoretical understanding of the construct/concept that is intended to be measured and the context in which it will be evaluated; 2) receiving input from a diverse expert panel (including potential instrument users) on the construct/concept; 3) assessing alignment of the instrument's content (including how to use it) with the construct/concept by the expert panel; and 4) using a framework to evaluate the correspondence between the instrument and the intended construct (201). We agree that content validity is a critical psychometric property since, without it, other measurement properties of the instrument will be less meaningful. That said, it is not a simple form of validity to establish or assess. One reason for this, as illustrated in our review, is that instruments measuring the same construct or attribute often differ regarding the context for which they were developed. For example, the IPC (64, 65) (attribute – *Evaluation Activity*) was developed to measure the perspectives of equity-deserving populations regarding interpersonal processes of care (e.g., timeliness,

ARTICLE IN PRESS

ARTICLE IN PRESS

communication, respectfulness) in acute/specialty care (65, 115, 116). However, multiple variations have been developed and are recommended for use in primary care (117-120), with potential for use in home/community care (114) and public health(64) (**Table 3**). Further illustrating this complexity, we found evidence regarding the comprehensibility of items when instruments are re-evaluated for content validity after translation into a different language or used for the first time in a different sector or setting. In these instances, our COSMIN ratings for instrument development quality were often 'inadequate' or 'doubtful' because the authors did not adequately describe their process. Other identified weaknesses in content validity for most of the 38 instruments across 169 psychometric papers involved authors: 1) not reporting the use of cognitive interviews to evaluate the comprehensiveness or comprehensibility of the instruments; 2) inadequate description of the content experts who appraised the instruments; and 3) not including intended users of the instruments in the content validation process. These shortcomings have been identified in recent systematic reviews (202-204) using COSMIN. In addition to

ARTICLE IN PRESS

ARTICLE IN PRESS

content validity, another dimension that could be used to assess the appropriateness of a tool is key informants' interpretation of the instrument's acceptability, relative advantage over existing measures, ease of completion, appropriateness to the context, fit with existing organizational activities, and the ability of the data garnered from the instrument to inform clinical or organizational decision making (43). From a Learning Health System perspective, engaging key informants in co-design will likely ensure instruments align with user needs, enhancing their relevance and effectiveness in improving health outcomes (205).

ARTICLE IN PRESS

Limitations

Although rigorous methods were used for this review, there are four main study limitations. First, we needed to utilize a more explorative and practical process to locate instruments by limiting our Stage 1 search to review articles, considering the complexity of ICON. Hence, while we identified 38 instruments, which were assessed in 169 papers across multiple healthcare sectors and settings, it is possible that not all relevant instruments were captured in these existing reviews.

ARTICLE IN PRESS

Similarly, instruments measuring concepts like the instruments we included may have been missed in our search, as they may not have been detailed in the articles (primarily reviews on context) we used to find tools. Further, we focused on instruments that assess a single context attribute at the meso- or macro-levels of ICON; hence, instruments that evaluate multiple attributes (n = 93 instruments) were not included in our analysis. Third, we primarily searched for psychometric articles in a single database (PubMed) using the methodological filter developed by Terwee and colleagues (54). While this may initially be perceived as a limitation, we found that the use of the precise methodological search filter in combination with reference checks was more sensitive than conducting a citation search(55). Fourth, although we analyzed our findings according to different sectors, which aligns with the idea that context can influence psychometric properties, the COSMIN manual (56, 57) does not currently provide guidance on analyzing data according to different sectors and settings.

ARTICLE IN PRESS

ARTICLE IN PRESS

Conclusion

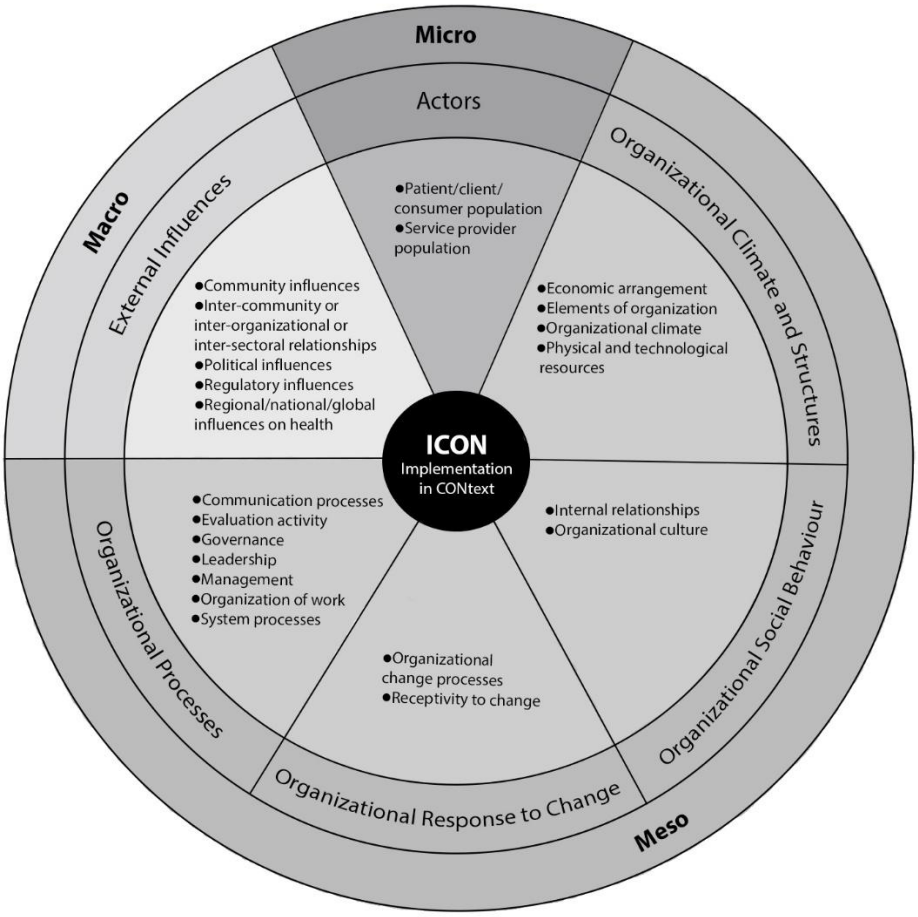
In summary, we located 38 instruments that measure ICON meso level attributes of context (25) and assessed 169 psychometric papers on these instruments based on the COSMIN guidelines (56, 57). Only 13 instruments met our criteria for recommending their use. We will be conducting a second psychometric review of instruments identified in the micro (individual) level of ICON and will then develop a comprehensive and pragmatic context survey that targets all 22 ICON attributes. Understanding and measuring context is crucial for the successful implementation and operation of Learning Health Systems, which rely on continuous improvement informed by contextual factors. Future work should also include: 1) additional psychometric assessments of the 25 instruments not yet recommended for use based on our analysis, given inadequate existing evidence, 2) psychometric and pragmatic assessments of instruments that assess multiple attributes at the meso- or macro levels of ICON; 3) the development of additional instruments to target the ICON attributes important for successful implementation and for which instruments were not identified (n=6

ARTICLE IN PRESS

meso-level and n=5 macro-level), and 4) the development of an open access registry of the instruments identified in this review and a website to facilitate access to the instruments and their psychometric and pragmatic properties.

ARTICLE IN PRESS

Figure 1. The **I**mplementation in **Con**text (ICON) Framework.



Outer circle represents levels of context. Middle circle represents the 6 context domains within ICON. Inner circle represents the 22 core attributes of context within each domain in ICON.

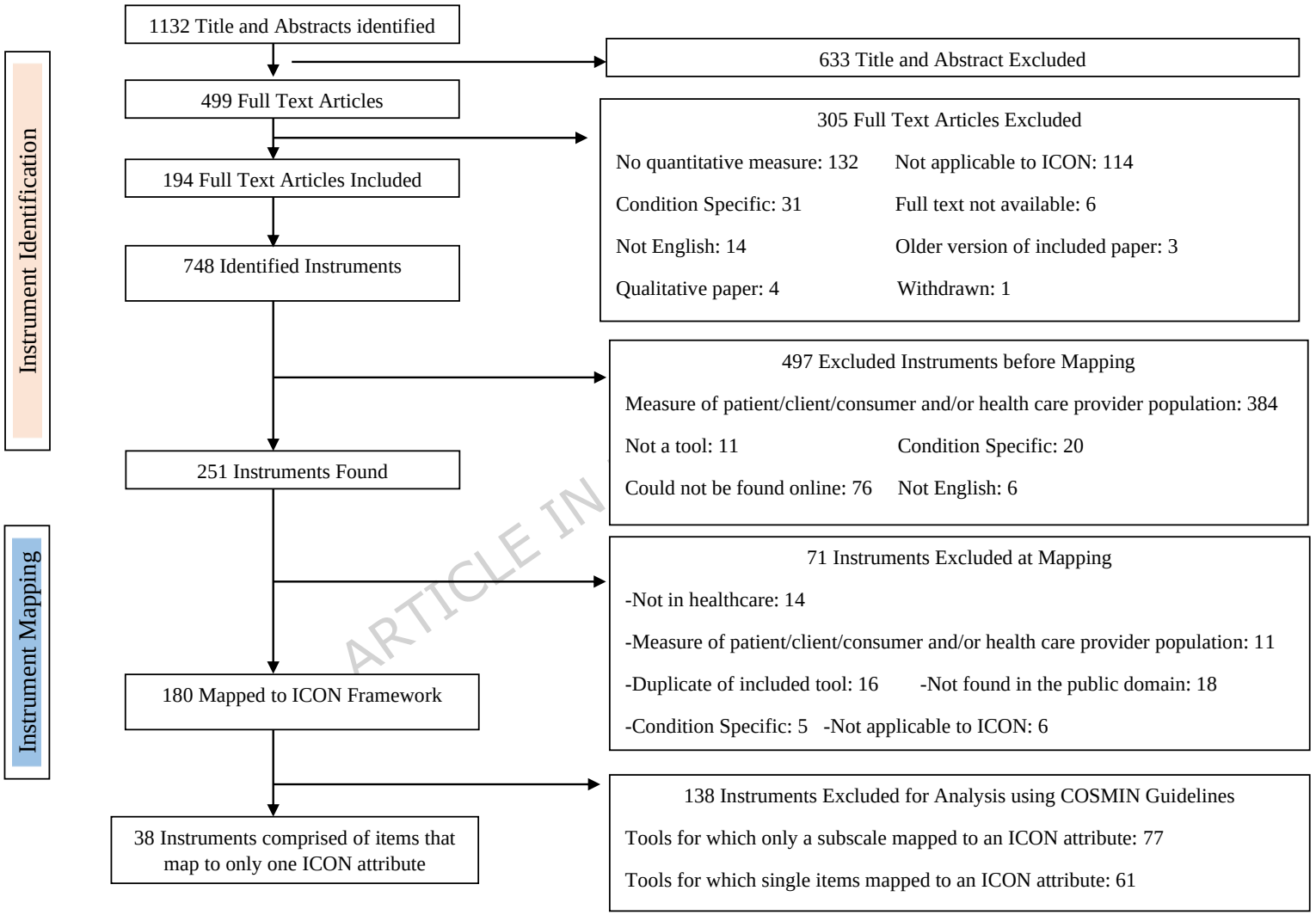
ARTICLE IN PRESS

Table 1. Eligibility Criteria

Stage 1 Search	Stage 2 Search
Inclusion criteria • Articles reporting studies that included an instrument measuring one or more ICON attribute(s)	Inclusion criteria • Articles reporting on the validity and reliability of the included instruments used in any healthcare setting and sector • Articles that described the development of instruments, regardless of the context (i.e., even if they did not originate in healthcare)
Exclusion criteria • Qualitative and case study designs • Study protocols, conference abstracts, editorials, opinion papers • Articles not published in English	Exclusion criteria • Synthesis and case study designs • Articles reporting studies that did not evaluate the instrument's psychometric properties • Study protocols, conference abstracts, editorials, opinion papers • Articles not published in English • Articles not available in full text • Articles reporting instruments for which permission and payment was required

ARTICLE IN PRESS

Figure 2. PRISMA Diagram – Identifying Tools Relevant to ICON



ARTICLE IN

ARTICLE IN PRESS

Figure 3. Number of Instruments Included according to ICON Attributes

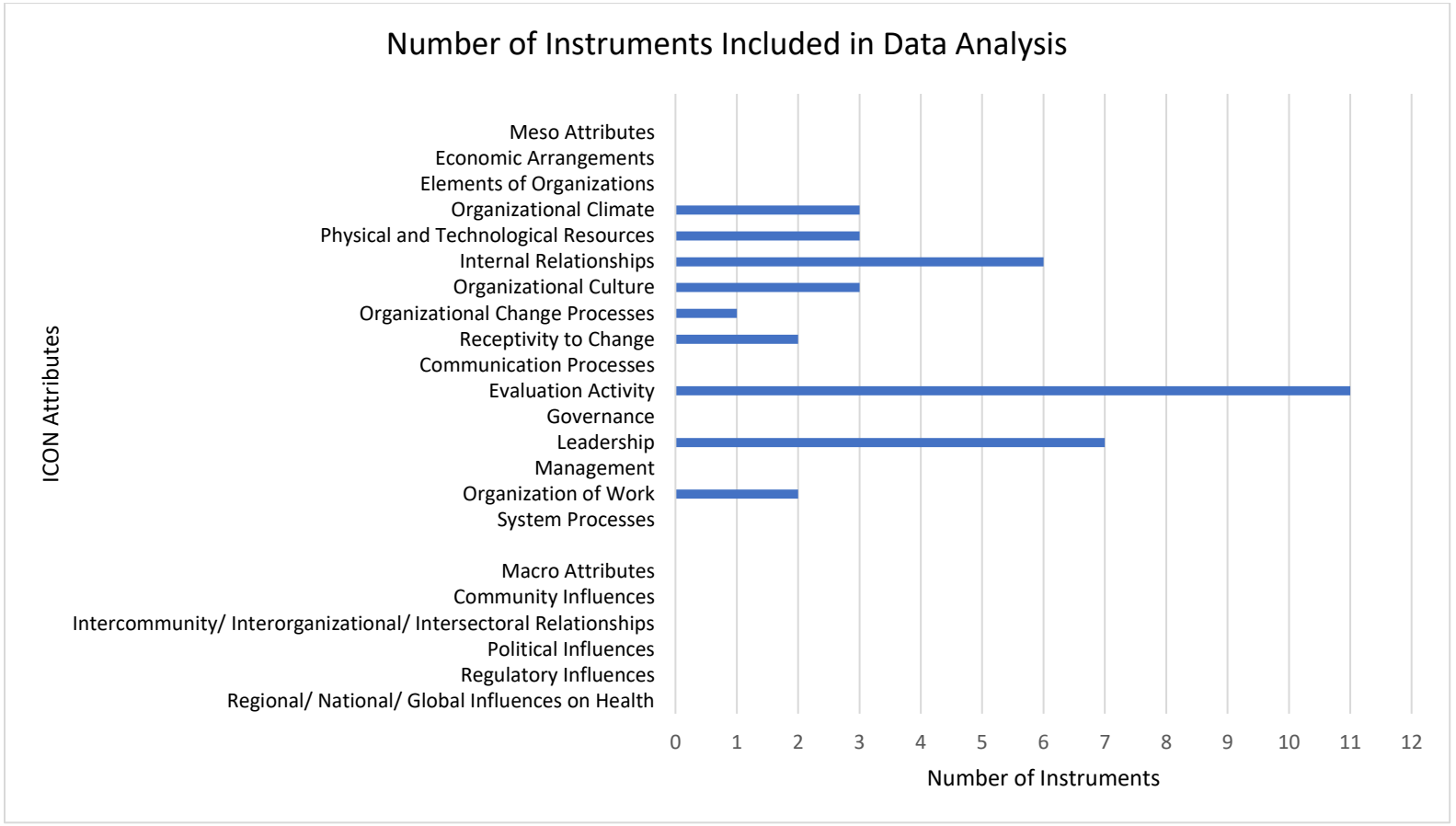
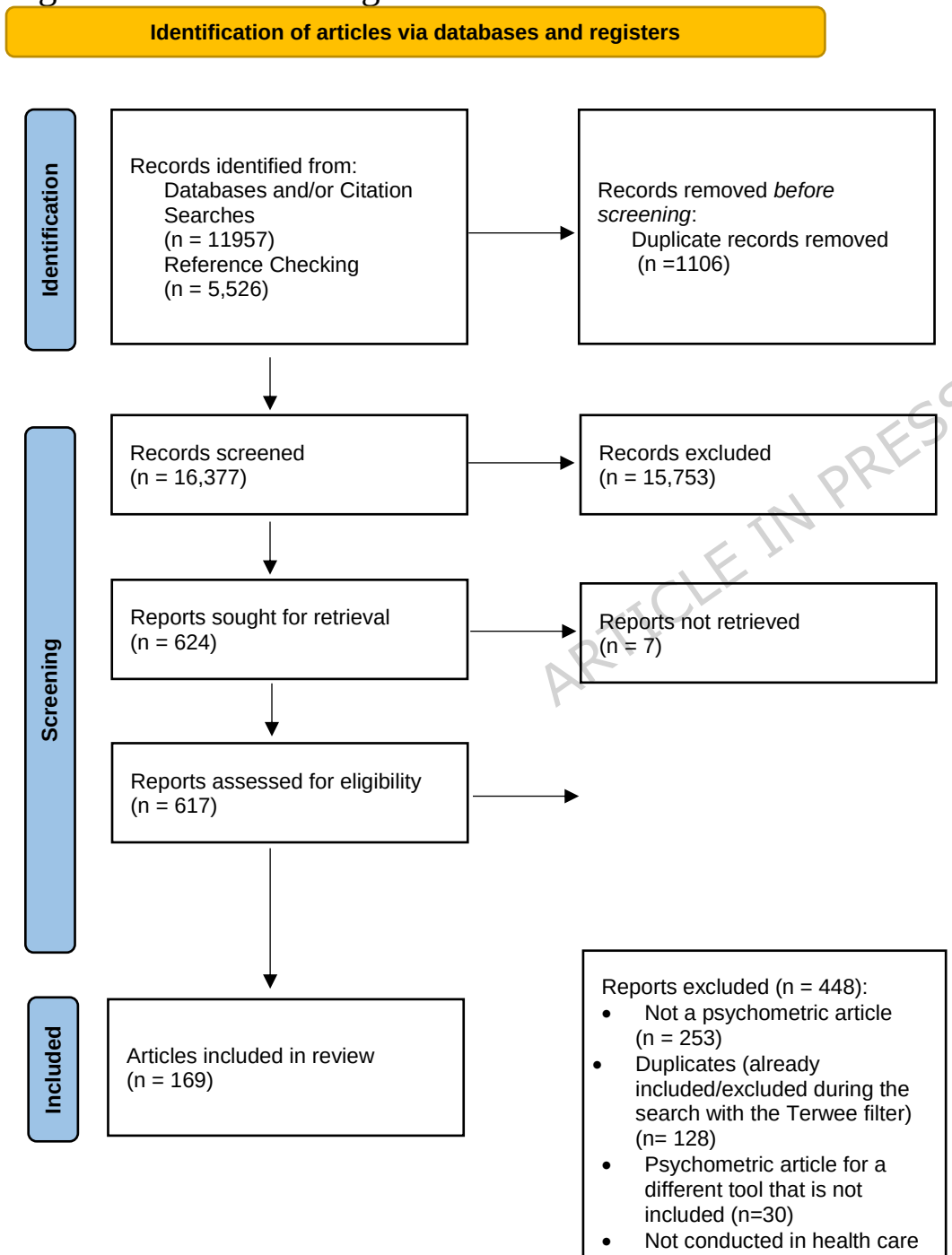


Figure 4. PRISMA Diagram



ARTICLE IN PRESS

ARTICLE IN PRESS

Table 2. Description of Instruments

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
Domain: Organizational Climate and Structures									
Organization Climate									
Costa and Anderson (2011)'s (82) measure of trust in teams	1(82)	Costa, 2011(82)	Team Climate	Sector: Acute or Specialty Care (82) Setting: Hospital and social care institutions (82)	Employees of hospitals and social care institutions (82)	English(82)	1) Propensity to trust 2) Perceived trustworthiness 3) Cooperative behaviours 4) Monitoring behaviours	21 (82)	7-point Likert scale (completely disagree to completely agree) (82)

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				Country: Netherlands (82)					
Internal Participation Scale	1(66)	Körner, 2013(66)	Team Climate	Sector: Rehabilitation(66) Setting: Inpatient medical rehabilitation clinics (66)	Service providers and patient or consumer population (66)	English (66)	Unidimensional scale (66)	6 (66)	4-point Likert scale (does not apply at all to fully applies) (66)

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				Country: Germany (66)					
Team Climate Inventory (TCI)	8 (62, 63, 98-101, 104, 206)	Anderson et al. (1996, 1998) (62, 63)	Team Climate	Sector: Acute or Specialty Care (62, 63, 98-100); Home or Community Care (206); Primary Care (101, 104)	Service providers (63, 99-101, 104, 206), quality improvement teams (98), Manage	English (62, 63, 100) (Dutch (98, 206), Portuguese(104), Finish (99),	1) Vision 2) Participatory safety 3) Support for innovation 4) Task orientation Interaction Frequency	Original: 38 (63) Short form: 14	Vision: 7-point Likert scale (1= not at all to 7 = completely) Participative safety: 5-point response

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				Setting: Hospital (62, 63, 98); Broader government and social sector(99, 100); Home-based elderly care (206); Primary care (101, 104)	ment (62)	French (101))		(206)	scale (1 = a very little extent to 5 = a very great extent) Task orientation: 5-point response scale (1 = strongly disagree

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				Country: Brazil (104), Canada (101), Finland (99), Netherlands (98, 206), UK (62, 63, 100)					to 5 = strongly agree) Support for innovation: 5-point response scale (1 = strongly disagree to 5 = strongly agree) (62, 63)

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
Physical and Technological Resources									
Questionnaire on Computer Systems and Decision Making	0	Hatcher, 1997 (207)	Technology and Equipment & Supplies	Sector: Acute or Specialty care (207) Setting: Hospitals (207) Country: United States of America (207)	Information system personnel at hospitals (207)	English (207)	1) Facility information 2) Information system approach 3) Information systems design 4) Impact of information systems 5) Face to face meetings	41 (207)	A mixture of open-ended questions, yes or no questions, and 5-point Likert scale (not at all to quite extensiv

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
							6) Decision making Internet Usage		ely) (207)
Murphy et al. (1994)'s (68) measure of nurses' perception of a patient care information system	1	Murphy, 1994 (68)	Technology and Equipment & Supplies	<u>Sector:</u> Acute or Specialty care (68) <u>Setting:</u> Hospitals (68) <u>Country:</u> United States of	Service providers (Nurses) (68)	English (68)	Unidimensional scale (68)	12 (68)	5-point Likert scale (description of scale was not reported) (68)

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				America (68)					
Perceived Health Web Site Usability Questionnaire (PHWSUQ)	2 (67, 105)	Nahm, 2006 (67)	Online Resources	Sector: Other (67, 105) Setting: Not identified (67) and inpatient and outpatient settings within a large integrated health care	Patient or consumer population (67, 105)	English (67, 105)	1) Satisfaction 2) Ease of Reading 3) Usefulness	10 (67, 105)	7-point Likert scale (very unsatisfied to very satisfied) (67)

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				system (105) <u>Country</u>: United States of America (67, 105)					
Domain: Organizational Social Behaviour									
Internal Relationships									
Collaborative Practice Scale	2 (69, 208)	Weiss, 1985 (69)	Partnerships (or Collaborative Practices)	<u>Sector</u>: Acute or Specialty care (69, 208)	Service providers (Nurses and Physicians) (69, 208)	English (69) (German (208))	1) Nurses 2) Physicians	19 (69, 208)	6-point Likert scale (always to never) (69, 208)

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article(s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				Setting: Hospitals (69, 208) Country: Switzerland (208); United States of America (69)					
Community Impacts of Research Oriented	1 (94)	King 2009 (94)	Partnerships (or Collaborative Practices)	Sector: Home or community (94)	Researchers and community partners (94)	English (94)	1) Personal knowledge development 2) Personal research	33 (94)	7-point Likert scale (not at all to a very

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
Partnerships (CIROP)				Setting: Community based health research (94) Country: Canada (94)			skill development 3) Organizational/group access to and use of information 4) Community and organizational development		great extent) (94)

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
Multiple-group measurement scale for interprofessional collaboration	2 (70, 209)	Kenaszchuk, 2010 (70)	Partnerships (or Collaborative Practices)	Sector: Acute or Specialty care (70, 209) Setting: Hospitals (70, 209) Country: Canada (210), Italy (209)	Service providers (70, 209)	English (70) (German (209), Italian (209))	1) Communication 2) Accommodation 3) Isolation	13 (70)	4-point Likert scale (strongly disagree to strongly agree) (70)
Shared Decision-Making	7 (71, 106,	Scholl, 2012 (71)	Partnerships (or Collabor	Sector: Acute or Specialty	Service providers	German (71, 138)	Unidimensional scale (71)	9 (71)	6-point Likert scale

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
Questionnaire -Doc (SDM-Q-Doc)	128, 129, 138, 211, 212)		ative Practices)	care (106, 128, 129, 211); Primary care (71, 138, 212) <u>Setting:</u> Hospital (106, 211); Oncology clinical settings (128); Specialty outpatient clinics	(Specifically physicians) (71, 106, 128, 129, 138, 211, 212)	(English (71, 138), Japanese (106, 212), Spanish (211), Persian (128), Dutch (129))			(completely disagree to completely agree) (71)

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article(s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				(e.g., psychiatry, oncology etc.) (129), Primary care clinics (71, 138, 212) <u>Country:</u> Iran (128), Japan (106, 212), Germany (71, 138), Spain (211),					

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				Netherlands (129)					
Team Decision Making Questionnaire	1 (96)	Batorowicz, 2008 (96)	Partnerships (or Collaborative Practices)	<u>Sector:</u> Acute or Speciality (96) <u>Setting:</u> Augmentative and Alternative Communication (AAC) Centres (96)	Service providers (96)	English (96)	Unidimensional scale (96)	19 (96)	7-point Likert scale (not at all to a vast extent) (96)

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				<u>Country:</u> Canada (96)					
Trust in Health Promotion Partnerships Scale	1(83)	Jones, 2011(83)	Partnerships (or Collaborative Practices)	<u>Sector:</u> Home or community care (83) <u>Setting:</u> Health promotion partnerships (hospitals, community	Service providers (83)	English (83)	Unidimensional scale (83)	14 (83)	5-point Likert scale (never to always) (83)

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				health services, voluntary groups etc.) (83) <u>Country:</u> Ireland (83)					
Organizational Culture									
On-line Patient Safety Climate Survey	1 (42)	Connelly, 2005 (42)	Cultural Norms	<u>Sector:</u> Acute or Speciality (42) <u>Setting:</u> Military	Service providers (42)	English (42)	1) Willingness to report errors 2) Problem solving processes	19 (42)	4-point Likert scale (strongly disagree to strongly

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				hospitals (42) <u>Country:</u> United States of America (42)			3) Attitude of leadership		agree) (42)
Stanford University Patient Safety Climate in Healthcare	1 (84)	Singer,2007 (84)	Cultural Norms	<u>Sector:</u> Acute or Speciality (84) <u>Setting:</u> Hospitals (84)	Service providers and management (84)	English (84)	1) Senior Managers' engagement 2) Organizational resources for safety	38 (84)	5-point Likert scale (never to always) (84)

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
Organizations questionnaire				<u>Country:</u> United States of America (84)			3) Overall emphasis on safety 4) Unit safety norms 5) Unit recognition and support for safety efforts 6) Fear of shame 7) Provision of safe care		

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
							8) Learning 9) Fear of blame		
Survey on Patient Safety in Ambulatory Care Organizations	1 (85)	Schutz, 2007 (85)	Cultural Norms	<u>Sector:</u> Acute or Speciality (85) <u>Setting:</u> Ambulatory clinics (pediatric, surgery, women's health, etc.) (85)	Service providers (85)	English (85)	1) Teamwork 2) Leadership 3) Communication 4) Tendency to Report Errors 5) Global Question	21 (85)	5-point Likert scale (strongly disagree to strongly agree) (85)

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				<u>Country:</u> United States of America (85)					
Domain: Organizational Response to Change									
Organizational Change Processes									
Howell and colleagues (72) measure of champion behaviour	1 (72)	Howell,2005 (72)	Champions/Opinion Leaders	<u>Sector:</u> Non-healthcare (72) <u>Setting:</u> Business	Champions of manufacturing firms (electronics, automoti	English (72)	1) Risk-taking propensity 2) Taking chances, innovativeness	15 (72)	5-point Likert scale (not at all to always) (72)

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				settings (72) <u>Country:</u> United States of America (72)	ve, and aerospace industries) (72)		3) Develops novel solutions to problems and achievement orientation		
Receptivity to Change									
Organizational Readiness for Implementing Change	6 (86, 108-112)	Shea, 2014 (86)	Readiness for Change	<u>Sector:</u> Acute care or specialty (86, 108, 109); Home or community	Service providers (86, 108-112)	English (86, 111, 112) (Danish (109),	1) Effectiveness 2) Commitment	12 (86)	5-point Likert scale (disagree to agree) (86)

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
(ORIC) scale				(111); Primary care (110, 112) <u>Setting:</u> Cancer services (108), Hospitals (86), Healthcare education (university) (86), Home care services		French (111), German (108), Portuguese (110))			

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article(s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				(111) Obstetrics (109), Primary care (110, 112) <u>Country:</u> Brazil (110), Canada (111), Denmark (109), Germany (108), South					

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				Africa (112), United States of America (86)					
Resistance to Change Scale	2 (73, 213)	Oreg, 2003 (73)	Compatibility and Readiness for Change	<u>Sector:</u> Other (73, 213) <u>Setting:</u> Universities (73, 213) <u>Country:</u> 17	Service providers (73, 213)	English (213) (Chinese, Croation, Czech, Germa	1) Routine seeking 2) Emotional reaction 3) Short-term focus 4) Cognitive rigidity	16 (73)	6-point Likert scale (strongly disagree to strongly agree) (73)

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				countries (73) (Australia, China, Croatia, the Czech Republic, Germany, Greece, Israel, Japan, Lithuania, Mexico, the Netherlands, Norway, Slovakia,		n, Greek, Hebrew, Japanese, Lithuanian, Spanish, Dutch, Norwegian, Slovakian, Spanish,			

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				Spain, Turkey, the United Kingdom, and the United States of America), United States of America (213)		Turkish) (73)			
Domain: Organizational Processes									
Evaluation Activity									

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
American Board of Internal Medicine's (ABIM) questionnaire	0	Webster, 1989 (59, 214)	Patient/Client/Consumer Feedback to Staff	<u>Sector:</u> Acute care or specialty (214) <u>Setting:</u> Hospitals (214) <u>Country</u>: United States of America (214)	Patient/Client/Consumer Population (214)	English (214)	Unidimensional	Complete: 23 (214) Short: 10 (214)	5-point Likert scale (poor to excellent) (214)
Care Transition	12 (97,	CTM-15:	Patient/Client/Con	<u>Sector:</u> Acute care	Patient/Client/Con	English (97,	1) Critical understand	CTM-	4-point Likert

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article(s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
s Measure (CTM)	173-181, 183, 215)	Coleman, 2002 (97) CTM-3: Parry, 2008 (176)	sumer Feedback to Staff	or specialty (173-181, 215); Primary care (97); Rehabilitation (183) <u>Setting:</u> Hospitals (173-178, 180, 181, 215), oncology centre (179); Primary	sumer Population (97, 176)	174, 176, 180, 183) (Arabic(179); Chinese (177, 178); Hebrew (179); Japanese (215); Korea	ing (6 items) 2) Preferences important (3 items) 3) Management preparation (4 items) 4) Care plan (2 items)	15: 15 (97) CTM-3: 3 (176)	scale (strongly disagree to strongly agree) (97, 176)

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				care clinics (97); Rehabilitation units (183) <u>Country:</u> Brazil (173); Canada (183); China (178); Israel (179); Japan		n (181); Swedish (175); Portuguese (173))			

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				(215); Korea (181); Singapore (177); Sweden (175); United States of America (97, 174, 176, 180)					
CollaboRATE measure	11 (95, 126,	Elwyn,2013 (95)	Patient/Client/Consumer	Sector: Acute care or specialty (95, 126,	Patient/Client/Consumer	English (95, 220, 223,	Unidimensional scale (95)	3	9-point Likert scale (no effort

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article(s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
	216-224)		Feedback to Staff	216-221); Home or community (222); Primary care (223); Unidentified (224) <u>Setting:</u> Ambulatory primary care clinics (223); Cancer research (219);	Population (95)	224) (Dutch (218); French (219); Spanish (126, 216, 221, 222); Swedish (217))			was made to every effort was made) (95)

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article(s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				CPAP clinics (217); Community mental health services (222); Hospitals (95, 216, 218, 221); Orthopaedic specialist clinics (220); Paediatric ambulatory					

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article(s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				clinics (126) <u>Country:</u> Argentina (216); Canary Islands (222); France (219); Sweden (217); Netherlands (218); United States of					

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				America (95, 126, 220, 221, 223, 224)					
Doctors' Interpersonal Skills Questionnaire (DISQ)	2 (74, 225)	Greco,1999 (74)	Patient/Client/Consumer Feedback to Staff	<u>Sector:</u> Primary care (74, 225) <u>Setting:</u> Primary care clinics (74, 225) <u>Country:</u> Australia (74);	Patient/Client/Consumer Population (74)	English (74, 225)	Unidimensional scale (74)	12	6-point Likert scale (varied with the item stem; e.g., not at all satisfied to extremely

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				United Kingdom (225)					satisfied) (74)
Interpersonal processes of care: IPC	9 (64, 65, 114-120)	Stewart, 2007 (65)	Patient/Client/Consumer Feedback to Staff	<u>Sector:</u> Acute or Specialty care (65, 115, 116); Primary care (117-120); Home or community (114); Public Health (64)	Patient/Client/Consumer Population (65)	English (64, 65, 114, 117-120) (French (118, 119); Japanese (116); Spanis	**Hurried Communication** 1) Lack of clarity 2) Hurried, distracted 3) Elicited concerns, responded 4) Explained results	IPC I = 41 (64) IPC II= 29 (65)	5-point Likert scale (never to always) (64, 65)

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				<u>Setting:</u> Adult Medical Clinic (public sector facility) (64) ; Diabetes specialty clinics (115); Hospital (65); Psychiatric clinics		h (114); Persian (115))	5) Explained medications **Decision Making** 6) Asked patient 7) Decided together **Interpersonal Style** 8) Emotional support, compassion 9) Respectfulness		

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article(s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				(116); Primary care clinics (117-120); Senior centers, community health clinics, unemployment agencies, and colleges (114)			10) Assumed socioeconomic status 11) Discriminated due to race/ethnicity 12) Disrespectful office staff		

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				<u>Country:</u> Canada (117-120); Iran (115); Japan (116); United States of America (64, 65, 114)					
Patient Participation Emergency	1 (87)	Frank 2011 (87)	Patient/Client/Consumer Feedback to Staff	<u>Sector:</u> Acute or Specialty care (87)	Patient/Client/Consumer Population (87)	Swedish (87)	1) Fight for participation 2) Requirement	17 (87)	4-point Likert scale (strongly disagree

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
Department questionnaire (PPED)				<u>Setting:</u> Hospital emergency departments (87) <u>Country:</u> Sweden (87)			for participation 3) Mutual participation 4) Participating in getting basic needs satisfied		to strongly agree) (87)
Risser Patient Satisfaction Scale	6 (88, 226-230)	Risser, 1975 (88)	Patient/Client/Consumer Feedback to Staff	<u>Sector:</u> Acute or Specialty care (226-230); Primary care (88)	Patient/Client/Consumer Population (88)	English (88, 228-230) (Greek (226, 227))	1) Technical-Professional 2) Interpersonal-educational	25 (88)	5-point Likert scale (strongly disagree to strongly

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				<u>Setting:</u> Hospitals (227, 229, 230); Oncology units (226); Orthopedic units (228); Primary care clinic (88) <u>Country:</u> Cyprus (226); Greece			3) Interpersonal-trusting		agree) (88)

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				(227); United States of America (88, 228-230)					
Role Based Performance Scale	1 (89)	Welbourne, 1998 (89)	Patient/Client/Consumer Feedback to Staff	<u>Sector:</u> Non-healthcare (89) <u>Setting:</u> Various industries, including traditional	Patient/Client/Consumer Population (89)	English (89)	1) Innovator 2) Career 3) Team 4) Organization 4) Job	20 (89)	5-point Likert scale (needs much improvement to excellent) (89)

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				manufacturing, service, and high technology (89) <u>Country:</u> United States of America (89)					
Shared Decision Making Questionn	21 (90, 121-	Simon, 2006 (90)	Patient/Client/Consumer	<u>Sector:</u> Acute or speciality care (90,	Patient/Client/Consumer	German (90, 127, 131,	Unidimensional scale (90)	9 (90)	6-point Likert scale (complet

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
aire (SDM-Q-9)	140, 231)		Feedback to Staff	121-135, 140); Primary care (90, 136-140) <u>Setting:</u> Breast cancer clinics (132, 133); Community outpatient behavioural heath (122);	Population (90)	138, 139) (Arabic (135); Chinese (132, 133); Dutch (129); English (122, 126, 132); Japanese (137);			ely disagree to completely agree) (90)

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				Hospitals (90, 123, 125, 129, 130); Multiple Sclerosis clinics (124, 127); Maternity care (131); Oncology units (128); Outpatient diabetic clinic (121); Paediatric		Hebrew (134); Hungarian (140); Malay (132); Persian (128); Romanian (123); Spanish (121, 122,			

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				clinics (126); Primary care (90, 136-140); Psychiatric clinics (134); Specialty clinics (135, 140) <u>Country:</u> Ecuador (121); Germany (90, 127,		124-126, 136))			

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				130, 131, 138, 139); Hungary (140); Iran (128); Israel (134); Japan (137); Malaysia (132); Netherlands (129); Romania (123); Spain (124, 125, 136);					

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				Taiwan (133); United States of America (122, 126); United Arab Emirates (135)					
The Patient Assessment of Chronic Illness	35 (91, 141-172, 232, 233)	Glasgow, 2005 (91)	Patient/Client/Consumer Feedback to Staff	<u>Sector:</u> Acute or speciality (141-151, 164, 166-169, 233);	Patient/Client/Consumer Population (91)	English (91, 142, 144, 154-156,	1) Patient activation: 2) Delivery system design/	PACIC = 20 (91)	5-point Likert scale (almost never to almost

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article(s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
Care Survey (PACIC)				Primary care (91, 152-161, 165, 166, 170-172, 232); Long-term care (166); Home or community (166); Public health (166); Rehabilitation (162);		162, 232, 233) (Arabic (167); Dutch (143, 164, 170); French (145, 164); German (164, 165); Italian	decision support 3) Goal setting/ tailoring 4) Problem solving / contextual 5) Follow up/ coordination	PACIC short = 11 (165) PACIC5A = 26 (172)	always) (91)

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				Unidentified (163) <u>Setting:</u> Cardiovascular disease care (143); Cardiac rehabilitation sites (162); Diabetes clinics (142, 144, 145, 149, 150, 167,		(164); Malay (152, 153); Portuguese (164); Spanish (141, 164))		PACIC with culturally sensitive items = 27	

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article(s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				169, 233); Heart transplant centres (164); Hospitals (141, 146-148, 151, 168); Primary care clinics (91, 153-161, 165, 170-172, 232); Varied settings				(232)	

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				(general practitioner office, hospices, public health, hospitals etc.) (166) <u>Country:</u> Australia (142, 159, 164); Belgium (164); Brazil (164);					

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				Canada (162-164); Denmark (149); Finland (150); France (148, 164); Germany (155, 164-166, 171, 172); Greece (168); India (147); Italy (164);					

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article(s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				Malaysia (152, 153); New Zealand (232); Netherlands (143, 151, 161, 170, 232); Saudi Arabia (167); Spain (164); Slovenia (160); South					

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				Korea (146); Switzerland (145, 164); Thailand (169); United Kingdom (156, 158, 164); United States of America (91, 141, 144, 154,					

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				157, 164, 233)					
Urgent Care System Questionnaire (UCSQ)	1 (92)	O'Cathain, 2011 (92)	Patient/Client/Consumer Feedback to Staff	<u>Sector:</u> Acute or speciality (92) <u>Setting:</u> Hospitals (92) <u>Country:</u> England (92)	Patient/Client/Consumer Population (92)	English (92)	1) Progress 2) Entry 3) Patient convenience	18 (92)	5-point Likert scale (strongly disagree to strongly agree) (92)
Leadership									

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
Ehrhart 2004 (75) Servant Leadership Scale	2 (75, 234)	Ehrhart, 2004 (75)	Leadership style	<u>Sector:</u> Home or Community (234); Non-healthcare <u>Setting:</u> Social Welfare Organizations (234); Grocery department stores (75)	Service providers (75)	English (75, 234)	Unidimensional scale (75)	14 (75)	5-point Likert scale (to a very small extent to a great extent) (75)

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				<u>Country:</u> Korea (234); United States of America (75)					
Global Transformational Leadership Scale	2 (76, 235)	Carless, 2000 (76)	Leadership style	<u>Sector:</u> Acute or Specialty care (235); Non-healthcare (76)	Service providers (76)	English (76, 235)	Unidimensional scale (76)	7 (76)	5-point Likert scale (rarely or never to very frequently, if not

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				<u>Setting:</u> Hospitals (235); Retail bank (76) <u>Country:</u> Pakistan (235); Australia (76)					always) (76)
Implementation Leadership Scale	9 (77, 184-189, 236, 237)	Aarons, 2014 (77)	Leadership style	<u>Sector:</u> Acute or Specialty care (77, 184-189); Home or	Service providers (77)	English (77, 184, 188, 189, 236,	1) Proactive leadership 2) Knowledgeable leadership	12 (77)	5-point Likert scale (not at all to very

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article(s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				community (236, 237) <u>Setting:</u> Alcohol and other drug treatment (184); Child welfare community-based organisation (236); Hospitals (185-188);		237) (Chinese (185, 186); Greek (187))	3) Supportive leadership 4) Perseverant leadership		great extent) (77)

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article(s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				Health services in schools (237); Mental health program (189); Outpatient mental health (77) <u>Country:</u> China (185, 186); Greece (187); United					

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article(s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				States of America (77, 184, 188, 189, 236, 237)					
Leadership Behavior Description Questionnaire	0	Stogdill, 1963 (238)	Leadership style	<u>Sector:</u> Non-health care (238) <u>Setting:</u> Diverse setting (commissioned and non-commissio	Service providers (238)	English (238)	1) Initiating Structure 2) Consideration	15 (238)	5-point Likert scale (never to always) (238)

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article(s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				ned police officers, executives in aircraft engineering staff, presidents of labour unions, presidents of colleges and universities, etc.) (238)					

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				<u>Country:</u> United States of America (238)					
Servant Leadership Questionnaire	1 (78)	Barbuto & Wheeler, 2006 (78)	Leadership style	<u>Sector:</u> Non-health care (78) <u>Setting:</u> Community leaders (78) <u>Country:</u> United	Service providers (78)	English (78)	1) Altruistic calling 2) Emotional healing 3) Wisdom 4) Persuasive mapping 5) Organizational stewardship	23 (78)	4-point Likert scale (strongly disagree to strongly agree) (78)

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				States of America (78)					
Servant leadership Survey	3 (79, 190, 191)	van Dierendonck, 2010 (79)	Leadership style	<u>Sector:</u> Non-health care (190); Not reported (79, 191) <u>Setting:</u> Private companies, non-profit organizations, or	Service providers (79)	Dutch (79) English (79) (Japanese (190); Spanish (191))	1) Empowerment 2) Accountability 3) Standing Back 4) Humility 5) Authenticity 6) Courage 7) Interpersonal	30 (79)	7-point Likert scale (strongly disagree to strongly agree) (79)

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				government agencies/public organizations (190); Not reported (79, 191) <u>Country:</u> Japan (190); Argentina (191); Spain (191);			acceptance 8) Stewardship		

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article(s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				Mexico (191); Netherlands (79); United Kingdom (79)					
The Mentorship Effectiveness Scale	1 (80)	Berk, 2005 (80)	Mentors	<u>Sector:</u> Healthcare education (80) <u>Setting:</u> Educational	Service providers (nursing faculty) (80)	English (80)	1) Description of Relationship 2) Outcome Measures	12 (80)	Open ended questions and select all that apply (80)

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				institutions (80) <u>Country:</u> United States of America (80)					
Organization of Work									
A107 NASA Task Load Index	6 (81, 192, 193, 239-241)	Hart and Stavelan d, 1988 (81)	Workloa d	<u>Sector:</u> Acute care or specialty (192, 193, 240); Healthcare Education	Service provider s (192, 193, 240)	Englis h (81, 192, 193, 239-241)	1) Mental Demand 2) Physical Demand 3) Temporal Demand 4) Performanc	6 (81)	7-point scale (very low to very high with 21

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article(s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				(239, 241) and Non-healthcare (Industry) (81) <u>Setting:</u> Intensive care departments (193, 240); Diabetes management (192); Healthcare education -			5) Effort 6) Frustration		gradations) (81)

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article(s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				University surgical training setting (239, 241); National Aeronautics and Space Administration (NASA) (81) <u>Country:</u> United States of					

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				America (81, 192, 193, 239, 240); Canada (241)					
RAFAELA System: 1) Professional Assessment of Nursing Care	10 (93, 194-199, 242-244)	Fagerstrom & Rainio, 1998 (93)	Workload	<u>Sector:</u> Acute or speciality (93, 194-198, 242-244); Home or community (199)	Service providers (93)	Finnish (93, 194-197, 242, 243) (Norwegian (198, 199),	PAONCIL: single item scale (93) OPC: 1) planning and coordination of nursing care	PAONCIL= 2 (93) OPC =	PAONCIL: Segmented line scale (very low to very high) (93)

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article (s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
Intensity Level (PAONCIL) and Oulu Patient Classification (OPC)				<u>Setting:</u> Hospitals (93, 194, 196, 242-244); Organ transplantation (198); Mixed speciality units from 8 hospitals (195); Gynaecology units (93); Home		Dutch (244))	2) breathing, blood circulation and symptoms of disease 3) nutrition and medication 4) personal hygiene and secretion 5) activity, sleep and rest 6) teaching, guidance in care and	6 item (195)	OPC: Nursing intensity classified as A (1 point), B (2 points), C (3 points), D (4 points); increasing points means increasin

ARTICLE IN PRESS

Instrument	# of articles	Reference of Development Article(s)	ICON Feature	Sector; Setting; Country	End User of Instrument	Original Language (translated languages)	Subscales	# of Items	Response Options
				healthcare units (199) <u>Country:</u> Finland (93, 194-197, 242, 243); Norway (198, 199); Amsterdam (244)			follow-up care emotional support (195)		g intensity (195)

ARTICLE IN PRESS

Table 3. Quality of the Instruments' Development Studies

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Domain: Organizational Climate and Structures												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Organization Climate												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Internal Participation	Very	Very	Very	Very	Very good	Doubtful	Doubtful	Very good	Inadequate	Doubtful	Inadequate	Inadequate

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Scale (66)	good	good	good	good								

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Team Climate Invento	Very	Very	Very	Very	Adequate	Doubtful	Doubtful	Very good	Doubtful	Doubtful	Doubtful	Doubtful

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
ry (TCI) (62, 63)	good	good	good	good								

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Costa and	Very	Very	Very	Very	Very good	Doubtful	Doubtful	Adequate	Very good	Very good	Doubtful	Doubtful

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Anderson (2011)'	good	good	good	good								

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
s (82) measure of												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
trust in teams												
Physical and Technological Resources												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Questionnaire on												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Computer System												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
s and Decision												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Making (58)												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Perceived Health	Very	Very	Very	Very	Very good	Inadequate	Inadequate	Inadequate	Inadequate	Doubtful	Inadequate	Inadequate

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Web Site Usabilit	good	good	good	good								

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
y Questionnaire												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
(PHWS UQ) (67)												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Murphy et al. (1994)’	Very	Very	Very	Very	Very good	Doubtful	Doubtful	Inadequate	Doubtful	Doubtful	Inadequate	Inadequate

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
s (68) measure of	good	good	good	good								

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
nurses' perception of a												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
patient care informa												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
tion system												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Domain: Organizational Social Behaviour												
Internal Relationships												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Collaborative Practic	Very	Very	Very	Very	Very good	Inadequate	Inadequate	Inadequate	Inadequate	Inadequate	Inadequate	Inadequate

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
e Scale (69)	good	good	good	good								

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Community Impacts	Very	Very	Very	Very	Very good	Very good	Very good	Very good	Very good	Very good	Very good	Very good

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
of Research	good	good	good	good								

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Oriented Partner												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
ships (CIROP) (94)												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Multiple-group measur	Very	Adequate	Very	Very	Adequate	Doubtful	Doubtful	Inadequate	Inadequate	Doubtful	Inadequate	Inadequate

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
ement scale for	good		good	good								

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
interprofessional												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
collaboration (70)												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Shared Decision-	Very	Very	Adequate	Very	Adequate	Doubtful	Doubtful	Inadequate	Inadequate	Doubtful	Inadequate	Inadequate

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Making Questionnaire -	good	good		good								

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Doc (SDM-												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Q-Doc) (71)												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Team Decision	Very	Very	Very	Very	Very good	Adequate	Adequate	Very good	Adequate	Adequate	Adequate	Adequate

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Making Questio	good	good	good	good								

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
nnaire (96)												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Trust in Health Promot	Very	Very	Very	Very	Very good	Adequate	Adequate	Very good	Doubtful	Very good	Doubtful	Doubtful

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
ion Partnerships	good	good	good	good								

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Scale (83)												
Organizational Culture												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
On-line Patient Safety	Very	Very	Very	Very	Very good	Doubtful	Doubtful	Very good	Doubtful	Doubtful	Doubtful	Doubtful

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Climate Survey (42)	good	good	good	good								

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Stanford	Very	Very	Very	Very	Very good	Doubtful	Doubtful	Doubtful	Doubtful	Doubtful	Doubtful	Doubtful

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
University Patient	good	good	good	good								

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Safety Climate in												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Healthcare Organi												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
zations questio												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
nnaire (84)												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Survey on Patient	Very	Very	Very	Very	Very good	Doubtful	Doubtful	Very good	Doubtful	Doubtful	Doubtful	Doubtful

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Safety in Ambul	good	good	good	good								

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
atory Care Organi												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
zations (85)												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Domain: Organizational Response to Change												
Organizational Change Processes												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Howell and colleag	Very	Very	Very	Very	Very good	Doubtful	Doubtful	Inadequate	Inadequate	Inadequate	Inadequate	Inadequate

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
ues (72) measur	good	good	good	good								

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
e of champion												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
behaviour												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Receptivity to Change												
Organizational	Very	Very	Very	Very	Very good	Doubtful	Doubtful	Adequate	Doubtful	Doubtful	Doubtful	Doubtful

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Readiness for Implem	good	good	good	good								

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
enting Change (ORIC)												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
scale (86)												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Resistance to Change	Very	Very	Very	Very	Very good	Doubtful	Doubtful	Inadequate	Inadequate	Inadequate	Inadequate	Inadequate

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Scale (73)	good	good	good	good								

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Domain: Organizational Processes												
Evaluation Activity												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
American Board												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
of Internal Medici												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
ne's (ABIM)												

ARTICLE IN PRESS

questionnaire (59, 60)												
Care Transitions Measure (CTM) (97)	Very good	Very good	Very good	Very good	Adequate	Adequate	Adequate	Very good	Adequate	Very good	Adequate	Adequate
CollaboRATE measure (95)	Very good	Very good	Very good	Very good	Very good	Very good	Very good	Very good	Very good	Very good	Very good	Very good
Doctors' Interpersonal Skills Questionnaire	Very good	Doubtful	Very good	Very good	Adequate	Doubtful	Doubtful	Inadequate	Doubtful	Doubtful	Doubtful	Inadequate

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
(DISQ) (74)												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Interpersonal process	Very	Very	Very	Very	Very good	Doubtful	Doubtful	Very good	Very good	Doubtful	Doubtful	Doubtful

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
es of care:	good	good	good	good								

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
IPC -29 (64, 65)												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Patient Participation	Very	Very	Very	Very	Very good	Doubtful	Doubtful	Very good	Doubtful	Doubtful	Doubtful	Doubtful

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Emergency Depart	good	good	good	good								

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
ment questionnaire												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
(PPED) (87)												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Risser Patient Satisfac	Very	Very	Very	Very	Very good	Doubtful	Doubtful	Very good	Doubtful	Doubtful	Doubtful	Doubtful

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
tion Scale (88)	good	good	good	good								

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Role Based	Very	Very	Very	Very	Very good	Doubtful	Doubtful	Very good	Doubtful	Doubtful	Doubtful	Doubtful

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Performance	good	good	good	good								

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Scale (89)												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Shared Decision	Very	Very	Very	Very	Very good	Doubtful	Doubtful	Very good	Doubtful	Doubtful	Doubtful	Doubtful

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Making Questionnaire	good	good	good	good								

ARTICLE IN PRESS

(SDM-Q-9) (90)												
The Patient Assessment of Chronic Illness Care Survey (91)	Very good	Very good	Very good	Very good	Very good	Doubtful	Doubtful	Very good	Doubtful	Doubtful	Doubtful	Doubtful
Urgent Care System Questionnaire (UCSQ) (92)	Very good	Very good	Very good	Very good	Very good	Doubtful	Doubtful	Very good	Doubtful	Doubtful	Doubtful	Doubtful
Leadership												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Ehrhart 2004 Servant	Very	Very	Very	Very	Adequate	Doubtful	Doubtful	Inadequate	Doubtful	Doubtful	Inadequate	Inadequate

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Leadership	good	good	good	good								

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Scale (75)												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Global Transformation	Very	Very	Very	Very	Very good	Doubtful	Doubtful	Inadequate	Doubtful	Doubtful	Inadequate	Inadequate

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
al Leadership	good	good	good	good								

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Scale (76)												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Implementation	Very	Very	Very	Very	Very good	Doubtful	Doubtful	Inadequate	Doubtful	Doubtful	Inadequate	Inadequate

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Leadership	good	good	good	good								

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Scale (77)												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Leadership Behavi												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
or Description												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Questionnaire (61)												

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
Servant Leadership	Very	Very	Very	Very	Adequate	Doubtful	Doubtful	Inadequate	Doubtful	Doubtful	Inadequate	Inadequate

ARTICLE IN PRESS

Questionnaire (78)	good	good	good	good								
Servant leadership Survey (79)	Very good	Very good	Very good	Very good	Very good	Doubtful	Doubtful	Inadequate	Doubtful	Doubtful	Inadequate	Inadequate
The Mentorship Effectiveness Scale (80)	Very good	Very good	Very good	Very good	Very good	Doubtful	Doubtful	Inadequate	Doubtful	Doubtful	Inadequate	Inadequate
Organization of Work												
NASA Task Load Index (81)	Very good	Very good	Very good	Very good	Very good	Doubtful	Doubtful	Inadequate	Doubtful	Doubtful	Inadequate	Inadequate

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
RAFAELA	Very	Very	Very	Very	Very good	Doubtful	Doubtful	Very good	Doubtful	Doubtful	Doubtful	Doubtful

ARTICLE IN PRESS

Instrument	Instrument design							Cognitive interview (CI) study or pilot test				
	General design requirements							General design requirements				
	Clear construct	Clear origin of construct	Clear target population for which the instrument was developed	Clear context of use	Instrument developed in sample representing the target population	Concept elicitation (relevance and comprehensiveness)	Total instrument design	CI study performed in sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total instrument development
System (93)	good	good	good	good								

ARTICLE IN PRESS

ARTICLE IN PRESS

Table 4. Measurement Property Rating and Adapted GRADE according to Sector Organized by Attributes

		Psychometric properties (Rating (number of pertinent articles) COSMIN adapted GRADE quality of evidence)									
Instrument	Sector	CV	SV	IC	CCV/ MI	ME	DCV	DKG	R	RL	COSMIN Rating
Domain: Organizational Climate and Structures											
Organizational Climate											
Costa and Anderson (2011)'s (82) measure of trust in teams	Acute care or specialty (n=1) (82)	+ (n=1); Low (82)	+ (n=1); High (82)	+ (n=1); High (82)			+ (n=1); High (82)			+ (n=1); High (82)	A
Internal Participation Scale	Rehabilitation (n=1) (66)	+ (n=1); Low (66)	+ (n=1); High (66)	+ (n=1); High (66)			+ (n=1); High (66)				A

ARTICLE IN PRESS

		Psychometric properties (Rating (number of pertinent articles) COSMIN adapted GRADE quality of evidence)									
Instrument	Sector	CV	SV	IC	CCV/ MI	ME	DCV	DKG	R	RL	COSMIN Rating
Team Climate Inventory (TCI)	Acute care or specialty (n=3) (62, 63, 98)	+ (n=1); Low (62)	+ (n=2); High (63, 98)	+ (n=3); High (62, 63, 98)				+ (n=2); Low (63, 98)		+ (n=1); Low (63)	A
	Primary Care (n=2) (101, 104)	+ (n=1); Low (104)	? (n=2); Moderate (101, 104)	+(n=2); High (101, 104)			+(n=1); Low (104)	+(n=1); Low (101)			A
	Home or community		+ (n=1);	+ (n=1);			+ (n=1); Low (206)		- (n=1);		B

ARTICLE

		Psychometric properties (Rating (number of pertinent articles) COSMIN adapted GRADE quality of evidence)									
Instrument	Sector	CV	SV	IC	CCV/ MI	ME	DCV	DKG	R	RL	COSMIN Rating
	(n=1) (206)		High (206)	High (206)					Low (206)		
	Public Health (n=2) (99, 100)		+ (n=2); High (99, 100)	+ (n=2); High (99, 100)			+ (n=1); High (99)	- (n=1); Low (99)			B
Physical and Technological Resources											
Questionnaire on Computer Systems and Decision	No psychometric paper										

ARTICLE IN PRESS

		Psychometric properties (Rating (number of pertinent articles) COSMIN adapted GRADE quality of evidence)									
Instrument	Sector	CV	SV	IC	CCV/ MI	ME	DCV	DKG	R	RL	COSMIN Rating
Making (58)											
A207 Murphy et al. (1994)'s measure of nurses' perception of a patient care information system	Acute or Specialty (n=1) (68)	+ (n=1); Low (68)	? (n=1); Low (68)	? (n=1) Low (68)				- (n=1); Low (68)			B
Perceived Health Web Site	Public Health (n=2)	? (n=1);	? (n=1);	? (n=2);			+ (n=1); Low (105)				B

ARTICLE

		Psychometric properties (Rating (number of pertinent articles) COSMIN adapted GRADE quality of evidence)									
Instrument	Sector	CV	SV	IC	CCV/ MI	ME	DCV	DKG	R	RL	COSMIN Rating
Usability Questionnaire (PHWSUQ)	(67, 105)	Low (67)	Low (105)	Moderate (67, 105)							
Domain: Organizational Social Behaviour											
Internal Relationships											
Collaborative Practice Scale	Acute or Specialty (n=2) (69, 208)		? (n=2); Moderate (69, 208)	? (n=2); High (69, 208)			+ (n=2); Moderate (69, 208)			+/- (n=1); Low (69)	B
Community Impacts of	Home or commun	+ (n=1);	? (n=1);	? (n=1);			+/- (n=1);	+ (n=1)		- (n=1); Low (94)	B

ARTICLE IN PRESS

		Psychometric properties (Rating (number of pertinent articles) COSMIN adapted GRADE quality of evidence)									
Instrument	Sector	CV	SV	IC	CCV/ MI	ME	DCV	DKG	R	RL	COSMIN Rating
Research Oriented Partnerships (CIROP)	ity (n=1) (94)	High (94)	Moderate (94)	High (94)			Low (94)	; Low (94)			
Multiple-group measurement scale for interprofessional collaboration	Acute or Specialty (n=1) (70)	? (n=1); Very low (70)	+ (n=1); High (70)	+ (n=1); High (70)			+ (n=1); High (70)			+ (n=1); Low (70)	B
	Public Health	? (n=1)	+ (n=1);	+ (n=1);							B

ARTICLE IN PRESS

		Psychometric properties (Rating (number of pertinent articles) COSMIN adapted GRADE quality of evidence)									
Instrument	Sector	CV	SV	IC	CCV/ MI	ME	DCV	DKG	R	RL	COSMIN Rating
	(n=1) (209)	; Low (209)	High (209)	High (209)							
Shared Decision-Making Questionnaire -Doc (SDM-Q-Doc)	Acute or Specialty (n=3) (128, 129, 211)	? (n=3); Low (128, 129, 211)	+ (n=3); High (128, 129, 211)	+ (n=3); High (128, 129, 211)			+ (n=3); Moderate (128, 129, 211)	+ (n=1); High (211)			B
	Primary Care (n=2) (138, 212)	? (n=1); Low (138, 212)	+ (n=1); High (212)	+ (n=1); High (212)	+ (n=1) Low (138)		+ (n=1) Moderate (212)			- (n=1); Low (138)	B
	Long Term Care	+ (n=2);	+ (n=2);	+ (n=2);	+ (n=1)					- (n=1); Low (71)	A

ARTICLE IN PRESS

		Psychometric properties (Rating (number of pertinent articles) COSMIN adapted GRADE quality of evidence)									
Instrument	Sector	CV	SV	IC	CCV/ MI	ME	DCV	DKG	R	RL	COSMIN Rating
	(n=2) (71, 106)	Very Low	Moderate (71, 106)	Moderate (71, 106)	; Low (106)						
Team Decision Making Questionnaire	Acute or specialty (n=1) (96)	+ (n=1); High (96)	? (n=1); High (96)	? (n=1); High (96)						+/- (n=1); Low (96)	B
Trust in Health Promotion Partnerships Scale	Public Health (n=1) (83)	+ (n=1); Low (83)	? (n=1); Moderate (83)	? (n=1); Low (83)							B
Organizational Culture											

ARTICLE

		Psychometric properties (Rating (number of pertinent articles) COSMIN adapted GRADE quality of evidence)									
Instrument	Sector	CV	SV	IC	CCV/ MI	ME	DCV	DKG	R	RL	COSMIN Rating
On-line Patient Safety Climate Survey	Acute or specialty (n=1) (42)	+ (n=1); Low (42)	? (n=1); Very low (42)	? (n=1); Very low (42)							B
Stanford University Patient Safety Climate in Healthcare Organizations questionnaire	Acute or specialty (n=1) (84)	+ (n=1); Low (84)	? (n=1); Moderate (84)	? (n=1); Low (84)							B

ARTICLE

		Psychometric properties (Rating (number of pertinent articles) COSMIN adapted GRADE quality of evidence)									
Instrument	Sector	CV	SV	IC	CCV/ MI	ME	DCV	DKG	R	RL	COSMIN Rating
Survey on Patient Safety in Ambulatory Care Organizations	Acute or specialty (n=1) (85)	+ (n=1); Low (85)									B
Domain: Organizational Response to Change											
Organizational Change Processes											
Howell and colleagues (72) measure of champion behaviour	Other-Business settings (n=1) (72)	+ (n=1); High (72)	+ (n=1); High (72)	+ (n=1); High (72)			+ (n=1); High (72)				A
Receptivity to Change											

ARTICLE

		Psychometric properties (Rating (number of pertinent articles) COSMIN adapted GRADE quality of evidence)									
Instrument	Sector	CV	SV	IC	CCV/ MI	ME	DCV	DKG	R	RL	COSMIN Rating
Organizational Readiness for Implementing Change (ORIC) scale	Acute or Specialty (n=2) (108, 109)	? (n=2); Moderate (108, 109)	+ (n=2); High (108, 109)	+ (n=2); Moderate (108, 109)							B
	Primary Care (n=2) (110, 112)	? (n=1); Low (110)	- (n=2); Moderate (110, 112)	+ (n=2); Moderate (110, 112)				? (n=1); Doubtful (112)		+ (n=1); Low (112)	B

ARTICLE

		Psychometric properties (Rating (number of pertinent articles) COSMIN adapted GRADE quality of evidence)									
Instrument	Sector	CV	SV	IC	CCV/ MI	ME	DCV	DKG	R	RL	COSMIN Rating
	Home or community (n=1) (111)	+ (n=1); Moderate (111)	- (n=1); High (111)	+ (n=1); High (111)							B
	Acute or Specialty, Public health, and non-health care (n=1) (86)	+ (n=1); Low (86)	+ (n=1); High (86)	+ (n=1); High (86)						+ (n=1); High (86)	A
Resistance to Change scale	Other-Universities	+ (n=1);	+ (n=2);	+ (n=2);	+ (n=1);		+ (n=2);	+ (n=1);			A

ARTICLE

		Psychometric properties (Rating (number of pertinent articles) COSMIN adapted GRADE quality of evidence)									
Instrument	Sector	CV	SV	IC	CCV/ MI	ME	DCV	DKG	R	RL	COSMIN Rating
	(n=2) (73, 213)	Low (73)	High (73, 213)	High (73, 213)	High (213)		High (73, 213)	High (73)			
Domain: Organizational Processes											
Evaluation Activity											
American Board of Internal Medicine's (ABIM) questionnaire (59, 60)	No psychometric paper										
Care Transitions Measure (CTM)	Acute or Specialty (n=10) (173-	+ (n=6); Low (173,	? (n=10);	+ (n=10); Mode			+ (n=5); Moderate (173, 177-	+ (n=7);		+ (n=3); Low (173,	A

ARTICLE IN PRESS

		Psychometric properties (Rating (number of pertinent articles) COSMIN adapted GRADE quality of evidence)									
Instrument	Sector	CV	SV	IC	CCV/ MI	ME	DCV	DKG	R	RL	COSMIN Rating
	181, 215)	175, 178, 179, 181, 215)	High (173-181, 215)	rate (173-181, 215)			179, 215)	Moderate (174, 176-178, 180, 181, 215)		178, 181)	
	Primary Care (n=1) (97)			+ (n=1); Low (97)			? (n=1); Low (97)			+ (n=1); Low (97)	B
	Rehabilitation (n=1) (183)	+ (n=1); Moderate (183)					+ (n=1); Very low (183)				B

		Psychometric properties (Rating (number of pertinent articles) COSMIN adapted GRADE quality of evidence)									
Instrument	Sector	CV	SV	IC	CCV/ MI	ME	DCV	DKG	R	RL	COSMIN Rating
CollaboRATE measure	Acute or Specialty (n=8) (95, 126, 216-221)	+ (n=5); Moderate (95, 126, 216, 219, 221)	? (n=2); Moderate (216, 217)	? (n=4); High (126, 216, 217, 219)			+ (n=6); Moderate (95, 126, 216-218, 220)	? (n=3); Moderate (219-221)		+ (n=1); Low (126)	B
	Primary care (n=1) (223)									+ (n=1); Mo(118)derate (223)	B
	Home or community (n=1) (222)		? (n=1);	? (n=1);				+ (n=1); Very			B

ARTICLE IN PRESS

		Psychometric properties (Rating (number of pertinent articles) COSMIN adapted GRADE quality of evidence)									
Instrument	Sector	CV	SV	IC	CCV/ MI	ME	DCV	DKG	R	RL	COSMIN Rating
			Moderate (222)	High (222)				low (222)			
	Unidentified (n=1) (224)						+ (n=1); High (224)	+ (n=1); Very low (224)	+ (n=1); Low (224)	+ (n=1); High (224)	B
Doctors' Interpersonal Skills Questionnaire (DISQ)	Primary care (n=2) (74, 225)	+ (n=1); Low (74)	? (n=2); Moderate (74, 225)	? (n=2); High (74, 225)			+ (n=1); High (74)	+ (n=1); Low (74)	+ (n=1); Very Low	+(1); Low (74)	B

		Psychometric properties (Rating (number of pertinent articles) COSMIN adapted GRADE quality of evidence)									
Instrument	Sector	CV	SV	IC	CCV/ MI	ME	DCV	DKG	R	RL	COSMIN Rating
									(74)		
Interpersonal processes of care (IPC)	Acute or Specialty (n=3) (65, 115, 116)	+ (n=2); Moderate (65, 115)	? (n=3); High (65, 115, 116)	+ (n=3); High (65, 115, 116)			+ (n=2); Moderate (115, 116)	+ (n=2); Very low (65, 115)		+ (n=2); Low (115, 116)	A
	Primary care (n=4) (117-120)	+ (n=1); Moderate (118)	? (n=2); Moderate (117, 120)	+ (n=2); Moderate (117, 120)	- (n=1); Low (119)		+ (n=1); High (117)				A
	Home or commun	+ (n=1)									B

ARTICLE IN PRESS

		Psychometric properties (Rating (number of pertinent articles) COSMIN adapted GRADE quality of evidence)									
Instrument	Sector	CV	SV	IC	CCV/ MI	ME	DCV	DKG	R	RL	COSMIN Rating
	ity (n=1) (114)	; Low (114)									
	Public health (n=1) (64)		? (n=1); Moderate (64)	? (n=1); High (64)							B
Patient Participation Emergency Department questionnaire (PPED)	Acute or Specialty (n=1) (87)	+ (n=1); Low (87)	? (n=1); Moderate (87)	? (n=3); High (87)			+ (n=1); Low (87)			+ (n=1); Low (87)	B

ARTICLE

		Psychometric properties (Rating (number of pertinent articles) COSMIN adapted GRADE quality of evidence)									
Instrument	Sector	CV	SV	IC	CCV/ MI	ME	DCV	DKG	R	RL	COSMIN Rating
Risser Patient Satisfaction Scale	Acute or Specialty (n=5) (226-230)	+ (n=3); Low (226, 227, 229)	? (n=3); Moderate (226, 229, 230)	? (n=5); Moderate (226-230)			+ (n=2); Very low (226, 227)	+ (n=1); Very low (228)		+ (n=2); Low (226, 227)	B
	Primary care (n=1) (88)	+ (n=1); Low (88)		? (n=1); High (88)						? (n=1); Low (88)	B
Role Based Performance Scale	Other - Non-healthcare (n=1) (89)	+ (n=1); Low (89)	- (n=1); Moderate (89)	? (n=1); High (89)			+ (n=1); High (89)				B

		Psychometric properties (Rating (number of pertinent articles) COSMIN adapted GRADE quality of evidence)									
Instrument	Sector	CV	SV	IC	CCV/ MI	ME	DCV	DKG	R	RL	COSMIN Rating
Shared Decision Making Questionnaire (SDM-Q-9)	Acute or Specialty (n=17) (90, 121-135, 140)	+ (n=11); Low (90, 121, 123, 126, 128, 129, 131, 132, 134, 135, 140)	+ (n=15); Moderate (90, 121-125, 128, 129, 131-135, 140)	+ (n=16); High (90, 121-126, 128-135, 140)	+ (n=2); Moderate (122, 128)		+ (n=9); Moderate (90, 125-131, 134)	+ (n=4); Low (125, 129, 131, 134)		+ (n=2); Very Low (128, 130)	A
	Primary care (n=6)	+ (n=5); Low	+ (n=6);	+ (n=5);			+ (n=2); Moderat	+ (n=2);		- (n=1); Low (138)	A

ARTICLE IN PRESS

		Psychometric properties (Rating (number of pertinent articles) COSMIN adapted GRADE quality of evidence)									
Instrument	Sector	CV	SV	IC	CCV/ MI	ME	DCV	DKG	R	RL	COSMIN Rating
	(90, 136-140)	(90, 136, 137, 139, 140)	Moderate (90, 136-140)	Moderate (90, 136, 137, 139, 140)			e (90, 137)	Low (137, 140)			
The Patient Assessment of Chronic Illness Care Survey (PACIC)	Acute or Specialty (n=16) (141-151, 164, 167-169, 233)	+ (n=9); Moderate (145, 147-151, 167-169)	? (n=15); High (141-146, 148-151, 164, 167-	+ (n=14); High (141, 143, 144, 146-151, 164, 167-			+ (n=6); Moderate (142, 146, 150, 151, 164, 168)	? (n=5); Low (144, 145, 148, 149, 168)	+ (n=1); High (143)	? (n=5); Moderate (141, 142, 144, 147, 169)	A

ARTICLE IN PRESS

		Psychometric properties (Rating (number of pertinent articles) COSMIN adapted GRADE quality of evidence)									
Instrument	Sector	CV	SV	IC	CCV/ MI	ME	DCV	DKG	R	RL	COSMIN Rating
			169, 233)	169, 233)							
	Primary care (n=16) (91, 152-161, 165, 170-172, 232)	+ (n=8); Low (91, 152, 153, 160, 161, 165, 171, 232)	? (n=15); Moderate (91, 152-161, 165, 170-172)	+ (n=15); High (91, 152-155, 157-161, 165, 170-172, 232)			+ (n=7); Moderate(91, 155, 158, 160, 161, 165, 171)	? (n=5); Moderate (153, 155, 170-172)		? (n=10); Moderate (91, 152, 153, 156, 158, 160, 161, 170-172)	A
	Rehabilitation			+ (n=1);			+ (n=1); Low (162)	+ (n=1);			B

ARTICLE IN PRESS

		Psychometric properties (Rating (number of pertinent articles) COSMIN adapted GRADE quality of evidence)									
Instrument	Sector	CV	SV	IC	CCV/ MI	ME	DCV	DKG	R	RL	COSMIN Rating
	(n=1) (162)			High (162)				Low (162)			
	Unidentified (n=1) (163)		+ (n=1); High (163)	+ (n=1); High (163)							B
	Acute or specialty; Primary care; Long-term care; Home or community; and	+ (n=1); Very low (166)	+ (n=1); Moderate (166)	+ (n=1); Low (166)							B

ARTICLE

		Psychometric properties (Rating (number of pertinent articles) COSMIN adapted GRADE quality of evidence)									
Instrument	Sector	CV	SV	IC	CCV/ MI	ME	DCV	DKG	R	RL	COSMIN Rating
	Public health (n=1) (166)										
Urgent Care System Questionnaire (UCSQ)	Acute or specialty (n=1) (92)	? (n=1); Low (92)	? (n=1); Moderate (92)	? (n=1); High (92)				+ (n=1); Low (92)			B
Leadership											
Ehrhart 2004 (75) Servant Leadership Scale	Home or Community (n=1) (234)						+ (n=1); Low (234)				B

ARTICLE

		Psychometric properties (Rating (number of pertinent articles) COSMIN adapted GRADE quality of evidence)									
Instrument	Sector	CV	SV	IC	CCV/ MI	ME	DCV	DKG	R	RL	COSMIN Rating
	Non-healthcare – (n=1) (75)		+ (n=1); Low (75)	+ (n=1); Low (75)			+ (n=1); High (75)				B
Global Transformational Leadership Scale	Acute or specialty (n=1) (235)		? (n=1); Low (235)	+ (n=1); High (235)			+ (n=1); High (235)	+ (n=1); High (235)			B
	Non-healthcare (n=1) (76)	? (n=1); Low (76)	+ (n=1); High (76)	+ (n=1); High (76)			+ (n=1); High (76)	+ (n=1); High (76)			B

ARTICLE IN PRESS

		Psychometric properties (Rating (number of pertinent articles) COSMIN adapted GRADE quality of evidence)									
Instrument	Sector	CV	SV	IC	CCV/ MI	ME	DCV	DKG	R	RL	COSMIN Rating
Implementation Leadership Scale	Acute or specialty (n=7) (77, 184-189)	+ (n=5); Low (77, 185-188)	+ (n=6); High (77, 184, 186-189)	+ (n=6); High (77, 184, 186-189)	+ (n=1); Moderate (188)		+ (n=5); High (77, 184, 186, 187, 189)			+ (n=4); Moderate (77, 184, 186, 188)	A
	Home or community (n=2) (236, 237)		+ (n=2); High (236, 237)	+ (n=2); High (236, 237)	+ (n=2); Low (236, 237)		+ (n=1); High (237)			+ (n=1); High (236)	B
Leadership Behavior Descriptio	No psychometric paper										

ARTICLE IN PRESS

		Psychometric properties (Rating (number of pertinent articles) COSMIN adapted GRADE quality of evidence)									
Instrument	Sector	CV	SV	IC	CCV/ MI	ME	DCV	DKG	R	RL	COSMIN Rating
n Questionnaire (61)											
Servant Leadership Questionnaire	Non-healthcare (n=1) (78)	+ (n=1); Low (78)	+ (n=1); High (78)	+ (n=1); High (78)			+ (n=1); Low (78)				A
Servant leadership Survey	Non-healthcare (n=3) (79, 190, 191)	+ (n=2); Low (79, 190)	+ (n=3); High (79, 190, 191)	+ (n=3); High (79, 190, 191)	+ (n=2); Moderate (79, 191)		+ (n=3); High (79, 190, 191)	+ (n=1); Low (191)		+ (n=1); High (190)	A

ARTICLE IN PRESS

		Psychometric properties (Rating (number of pertinent articles) COSMIN adapted GRADE quality of evidence)									
Instrument	Sector	CV	SV	IC	CCV/ MI	ME	DCV	DKG	R	RL	COSMIN Rating
The Mentorship Effectiveness Scale	Healthcare education (n=1) (80)	? (n=1); Very Low (80)									B
Organization of Work											
NASA Task Load Index	Acute or specialty (n=3) (192, 193, 240)		+ (n=2); High (192, 193)	+ (n=3); High (192, 193, 240)			+ (n=2); Moderate (192, 240)	+ (n=1); Low (240)			B
	Healthcare educatio	? (n=1);	? (n=1);	? (n=1);				? (n=2);			B

ARTICLE

		Psychometric properties (Rating (number of pertinent articles) COSMIN adapted GRADE quality of evidence)									
Instrument	Sector	CV	SV	IC	CCV/ MI	ME	DCV	DKG	R	RL	COSMIN Rating
	n (n=2) (239, 241)	Low (239)	Moderate (241)	High (241)				Very Low (239, 241)			
	Non-healthcare (n=1) (81)	+ (n=1); Low (81)	? (n=1); Very low (81)							+ (n=1); Very Low (81)	B
RAFAEL A System: Professional Assessment of Nursing Care	Acute or specialty (n=9) (93, 194-198, 242-244)	? (n=5); Low (93, 194, 198,		? (n=1); High (194)			+ (n=2); Low (93, 197)	+ (n=3); Low (196-198)		? (n=3); Very Low (194, 242, 244)	B

		Psychometric properties (Rating (number of pertinent articles) COSMIN adapted GRADE quality of evidence)									
Instrument	Sector	CV	SV	IC	CCV/ MI	ME	DCV	DKG	R	RL	COSMIN Rating
Intensity Level (PAONCIL) And Oulu Patient Classification instrument (OPCq)		243, 244)									
	Home or community (n=1) (199)			? (n=1); High (199)						- (n=1); Low(199)	B

Note: The measurement property ratings are rated as either: Sufficient (+); - Insufficient (-); Inconsistent (+/-); Not determined (?) and accompanied by the number of studies that evaluated the measurement property and an adapted GRADE measure of quality of evidence in brackets. The *n*

values represent the number of studies for which evaluated a particular psychometric property. CV= content validity; SV= structural validity; IC=Internal Consistency; CCV/ MI= Cross-cultural validity/Measurement Invariance; ME=Measurement Error; DCV= Discriminant/Convergent Validity; DKG= Discriminative/Known Groups Validity; R= Responsiveness; RL= Reliability. Blank cells mean they were not assessed. [1]- Two studies (90, 140) evaluated the SDM-Q in both acute/specialty care and primary care settings. There are also three studies that investigated both the SDM-Q and the SDM-Q-Doc (128, 129, 138).

ARTICLE IN PRESS

ARTICLE IN PRESS

List of Abbreviations

1. American Board of Internal Medicine's questionnaire = ABIM
2. Care Transitions Measure = CTM
3. Community Impacts of Research Oriented Partnerships = CIROP
4. COnsensus-based Standards for the selection of health Measurement Instruments = COSMIN
5. Doctors' Interpersonal Skills Questionnaire = DISQ
6. Implementation in Context = ICON
7. Interpersonal Processes of Care = IPC
8. Methodological Expectations of Cochrane Intervention Reviews = MECIR
9. Organizational Readiness for Implementing Change = ORIC
10. Oulu Patient Classification instrument = OPC
11. Outcomes Measurement Properties = OMP
12. Patient Assessment of Chronic Illness Care Survey = PACIC
13. Patient Participation Emergency Department questionnaire = PPED
14. Patient Reported Outcomes Measures = PROMs
15. Perceived Health Web Site Usability Questionnaire = PHWSUQ
16. Professional Assessment of Nursing Care Intensity Level = PAONCIL
17. Shared Decision-Making Questionnaire = SDM-Q-9
18. Shared Decision-Making Questionnaire -Doc = SDM-Q-Doc
19. Team Climate Inventory = TCI
20. Urgent Care System Questionnaire = UCSQ

ARTICLE IN PRESS

ARTICLE IN PRESS

Declarations

Ethics approval and consent to participate: Not applicable

Consent for publication: Not applicable

Availability of data and materials: The search strategy, the list of excluded articles, and pragmatic assessments are provided as additional files. The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests: The authors declare that they have no competing interests. Janet E Squires is an Associate Editor of *Implementation Science*. Alison M Hutchinson is the Protocols Editor for *Implementation Science* and *Implementation Science Communications*. They were not involved in the peer-review of this manuscript.

Funding: Canadian Institutes for Health Research; Grant

number: 426122

ARTICLE IN PRESS

Authors' Contributions: All the authors participated in the conception of the study. All authors contributed to the development of the study design. TR developed and ran the search strategy used to identify instruments included in the review (Stage 1). DS, JZ, LA, MDV, and WJS conducted screening and data extraction for Stage 1. AMH, IDG, JES, JZ, LA, LKC, MDV, and WJS mapped each individual item of an included instrument to an ICON framework. WJS ran the search for Stage 2 (psychometric review of included instruments). LKC, MG, VJW, and WJS conducted screening and data extraction for Stage 2. AMH, IDG, JES, LKC, MG, VJW, and WJS completed synthesis of the COSMIN and pragmatic assessments. AMH, IDG, JES, WJS drafted the manuscript. All team members critically reviewed and revised the manuscript regarding its content, approved the version that is to be published and agree to be accountable for all aspects of the work.

ARTICLE IN PRESS

Acknowledgements: We would like to thank Danielle Starr (DS) and Junqiang Zhao (JZ) for their contribution in the early phase of data collection. The ICON team: Chantal Backman[1], Anna Bergström[2], Jamie Brehaut[3,4], Melissa Brouwers[4], Christopher Burton[5], Christine Cassidy[6], Cheyne Chalmers[7], Anna Chapman[7,8], Heather Colquhoun[9], Janet Curran[6], Melissa Demery Varin[1], Paula Doering[10], Annette Elliott Rose[11], Christina Godfrey[12], Doris Grinspun[13], Gillian Harvey[14], Noah Ivers[15,16], Shelly-Anne Li[17], Letitia Nadalin-Penno[18], Tamara Rader[19], Jo Rycroft-Malone[20], Dawn Stacey[1,3], Sharon Straus[21,22,23], Andrea C. Tricco[21, 24], Lars Wallin[25]

[1] University of Ottawa, School of Nursing, Ottawa, Canada

[2] Uppsala University, Department of Women's and Children's Health, Uppsala, Sweden

[3] Ottawa Hospital Research Institute, Methods and Implementation Research Program, Ottawa, Canada

[4] University of Ottawa, School of Epidemiology and Public Health, Ottawa, Canada

[5] University of East Anglia, School of Health Science, Norwich, United Kingdom

[6] Dalhousie University, School of Nursing, Halifax, Canada

[7] Deakin University, School of Nursing and Midwifery, Geelong, Australia

[8] Deakin University, Centre for Quality and Patient Safety Research, Institute for Health Transformation, Geelong, Australia

ARTICLE IN PRESS

ARTICLE IN PRESS

[9] University of Toronto, Department of Occupational Science and Occupational Therapy, Toronto, Canada

[10] Élisabeth Bruyère Health, Ottawa, Canada

[11] Nova Scotia Health Authority, Nova Scotia, Canada

[12] Queen's University, School of Nursing, Kingston, Canada

[13] Registered Nurses' Association of Ontario (RNAO), Toronto, Canada

[14] Flinders University, College of Nursing and Health Sciences, Bedford Park
South Australia, Australia

[15] University of Toronto, Department of Family and Community Medicine, Toronto, Canada

[16] Women's College Hospital Research and Innovation Institute, Toronto, Canada

[17] Artists' Health Centre, Department of Family and Community Medicine, University Health Network, Toronto, Canada

[18] Faculty of Environmental Studies, Biotechnology and Health Sciences,
Canadore College North Bay, Toronto

[19] Canadian Agency for Drugs and Technologies in Health, Ottawa, Canada

[20] Lancaster University, Faculty of Health and Medicine, Lancaster, United Kingdom

ARTICLE IN PRESS

ARTICLE IN PRESS

[21] Li Ka Shing Knowledge Institute, St. Michael's Hospital, Unity Health Toronto, Ontario, Canada

[22] University of Toronto, Department of Medicine, Toronto, Ontario, Canada

[23] University of Toronto, Dalla Lana School of Public Health, Toronto, Ontario, Canada

[24] Epidemiology Division & Institute of Health Policy, Management, and Evaluation, Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada

[25] Dalarna University, School of Health and Welfare, Falun, Sweden

ARTICLE IN PRESS

ARTICLE IN PRESS

References

1. Grinspun D, Mazurek Melnyk B, Fineout-Overholt E, Naik S, Wallace K. Advancing optimal care with robust clinical practice guidelines. Evidence-Based Practice in Nursing & Healthcare: A Guide to Best Practice. 5th ed2023.
2. Grol R. Successes and failures in the implementation of evidence-based guidelines for clinical practice. Med Care. 2001;39(8 Suppl 2):II46–54.
3. Lang ES, Wyer PC, Haynes RB. Knowledge translation: closing the evidence-to-practice gap. Ann Emerg Med. 2007;49(3):355–63.
4. Lauer MS, Skarlatos S. Translational research for cardiovascular diseases at the National Heart, Lung, and Blood Institute: moving from bench to bedside and from bedside to community. Circulation. 2010;121(7):929–33.
5. McGlynn EA, SM A, Adams J, Keesey J, Hicks J, DeCristofaro A, et al. The quality of health care delivered to adults in the United States. N Engl J Med. 2003;348(26):2635–45.
6. Schuster MA, McGlynn EA, Brook RH. How good is the quality of health care in the United States? 1998. Milbank Q. 2005;83(4):843–95.
7. Runciman WB, Hunt TD, Hannaford NA, Hibbert PD, Westbrook JI, Coiera EW, et al. CareTrack: assessing the appropriateness of health care delivery in Australia. Med J Aust. 2012;197(2):100–5.
8. Seddon ME, Marshall MN, Campbell SM, Roland MO. Systematic review of studies of quality of clinical care in general practice in the UK, Australia and New Zealand. *Qual Health Care*. 2001;10(3):152–8.
9. Squires JE, Cho-Young D, Aloisio LD, Bell R, Bornstein S, Brien SE, et al. Inappropriate use of clinical practices in Canada: a systematic review. *CMAJ*. 2022;194(8):E279–E96.
10. Pomare C, Mahmoud Z, Vedovi A, Ellis LA, Knaggs G, Smith CL, et al. Learning health systems: a review of key topic areas and bibliometric trends. Learn Health Syst. 2022;6(1):e10265.
11. Lewis CC, Fischer S, Weiner BJ, Stanick C, Kim M, Martinez RG. Outcomes for implementation science: an enhanced systematic review of

ARTICLE IN PRESS

instruments using evidence-based rating criteria. Implement Sci. 2015;10(1):155.
12. Reichenpfader U, Carlfjord S, Nilsen P. Leadership in evidence-based practice: a systematic review. Leadersh in Health Serv. 2015;28(4):298–316.
13. Wozney L, McGrath PJ, Gehring ND, Bennett K, Huguet A, Hartling L, et al. eMental healthcare technologies for anxiety and depression in childhood and adolescence: Systematic review of studies reporting implementation outcomes. JMIR Ment Health. 2018;5(2).
14. Li S-A, Jeffs L, Barwick M, Stevens B. Organizational contextual features that influence the implementation of evidence-based practices across healthcare settings: a systematic integrative review. Syst Rev. 2018;7(1):72.
15. De Vries H, Bekkers V, Tummers L. Innovation in the public sector: a systematic review and future research agenda. Public Adm. 2016;94(1):146–66.
16. Lau R, Stevenson F, Ong BN, Dziedzic K, Treweek S, Eldridge S, et al. Achieving change in primary care—causes of the evidence to practice gap: systematic reviews of reviews. Implement Sci. 2015;11(1):40.
17. May C, Finch T, Mair F, Ballini L, Dowrick C, Eccles M, et al. Understanding the implementation of complex interventions in health care: the normalization process model. BMC Health Serv Res. 2007;7:148.
18. Coles E, Wells M, Maxwell M, Harris FM, Anderson J, Gray NM, et al. The influence of contextual factors on healthcare quality improvement initiatives: what works, for whom and in what setting? Protocol for a realist review. Syst Rev. 2017;6(1):168.
19. Taylor SL, Dy S, Foy R, Hempel S, McDonald KM, Øvretveit J, et al. What context features might be important determinants of the effectiveness of patient safety practice interventions? BMJ Qual Saf. 2011;20(7):611–7.
20. Bird M, MacPhee M, Shaw J, Wodchis WP, Jeffs L, Austin T, et al. Evaluating for learning and sustainability (ELS) framework: a realist synthesis. BMC Health Serv Res. 2025;25(1):683.

ARTICLE IN PRESS

ARTICLE IN PRESS

21. Squires JE, Graham I, Bashir K, Nadalin-Penno L, Lavis J, Francis J, et al. Understanding context: A concept analysis. J Adv Nurs. 2019;75(12):3448–70.
22. Squires JE, Graham ID, Hutchinson AM, Linklater S, Brehaut JC, Curran J, et al. Understanding context in knowledge translation: a concept analysis study protocol. J Adv Nurs. 2015;71(5):1146–55.
23. Squires JE, Aloisio LD, Grimshaw JM, Bashir K, Dorrance K, Coughlin M, et al. Attributes of context relevant to healthcare professionals' use of research evidence in clinical practice: a multi-study analysis. Implement Sci. 2019;14(1):52.
24. Squires JE, Hutchinson AM, Coughlin M, Bashir K, Curran J, Grimshaw JM, et al. Stakeholder Perspectives of Attributes and Features of Context Relevant to Knowledge Translation in Health Settings: A Multi-Country Analysis. International Journal of Health Policy and Management. 2022;11(8):1373–90.
25. Squires JE, Graham ID, Santos WJ, Hutchinson AM. The Implementation in Context (ICON) Framework: A meta-framework of context domains, attributes and features in healthcare. Health Res Policy Syst. 2023;21(1):81.
26. Packiam P, Graham ID, Graham N, Hutchinson A, Santos WJ, Squires JE. Assessing the applicability of the Implementation in CONtext (ICON) framework to low and middle income countries: A secondary analysis (In Review). Health Res Policy Syst. 2026.
27. Harvey G, Kitson A. PARIHS revisited: from heuristic to integrated framework for the successful implementation of knowledge into practice. *Implement Sci*. 2015;11(1):1–13.
28. Kitson A, Harvey G, McCormack B. Enabling the implementation of evidence based practice: a conceptual framework. Quality in Health Care. 1998;7(3):149–58.
29. Graham ID, Logan J, Harrison MB, Straus SE, Tetroe J, Caswell W, et al. Lost in knowledge translation: time for a map? J Contin Educ Health Prof. 2006;26(1):13–24.
30. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research

ARTICLE IN PRESS

findings into practice. A consolidated framework for advancing implementation science. Implement Sci. 2009;4:50.
31. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated Consolidated Framework for Implementation Research based on user feedback. Implement Sci. 2022;17(1):1–16.
32. Squires JE, Hutchinson AM, Coughlin M, Bashir K, Curran J, Grimshaw JM, et al. Stakeholder Perspectives of Attributes and Features of Context Relevant to Knowledge Translation in Health Settings: A Multi-Country Analysis. Int J Health Policy 2022;11(8):1373–90.
33. Chaudoir SR, Dugan AG, Barr CH. Measuring factors affecting implementation of health innovations: a systematic review of structural, organizational, provider, patient, and innovation level measures. Implement Sci. 2013;8:22.
34. Brennan SE, Bosch M, Buchan H, Green SE. Measuring organizational and individual factors thought to influence the success of quality improvement in primary care: a systematic review of instruments. Implement Sci. 2012;7(1):121.
35. Rogers L, De Brún A, McAuliffe E. Defining and assessing context in healthcare implementation studies: a systematic review. BMC Health Serv Res. 2020;20(1):591.
36. Society for Implementation Research Collaboration (SIRC). The SIRC Instrument Review Project (IRP): Society for Implementation Research Collaboration (SIRC),; 2020 [Available from: https://societyforimplementationresearchcollaboration.org/sirc-instrument-project/.
37. McHugh S, Dorsey CN, Mettert K, Purtle J, Bruns E, Lewis CC. Measures of outer setting constructs for implementation research: A systematic review and analysis of psychometric quality. Implement Res Pract. 2020;1:2633489520940022.
38. Powell BJ, Mettert KD, Dorsey CN, Weiner BJ, Stanick CF, Lengnick-Hall R, et al. Measures of organizational culture, organizational climate, and implementation climate in behavioral health: A systematic review. Implement Res Pract. 2021;2:26334895211018862.

ARTICLE IN PRESS

39. Dorsey CN, Mettert KD, Puspitasari AJ, Damschroder LJ, Lewis CC. A systematic review of measures of implementation players and processes: Summarizing the dearth of psychometric evidence. Implement Res Pract. 2021;2:26334895211002474.
40. Weiner BJ, Mettert KD, Dorsey CN, Nolen EA, Stanick C, Powell BJ, et al. Measuring readiness for implementation: A systematic review of measures' psychometric and pragmatic properties. Implement Res Pract. 2020;1:2633489520933896.
41. Stanick C, Halko H, Mettert K, Dorsey C, Moullin J, Weiner B, et al. Measuring characteristics of individuals: An updated systematic review of instruments' psychometric properties. Implement Res Pract. 2021;2:26334895211000458.
42. Connelly LM, Powers JL. Advances in Patient Safety
Online Patient Safety Climate Survey: Tool Development and Lessons Learned. In: Henriksen K, Battles JB, Marks ES, Lewin DI, editors. Advances in Patient Safety: From Research to Implementation (Volume 4: Programs, Tools, and Products). Rockville (MD): Agency for Healthcare Research and Quality (US); 2005.
43. Stanick CF, Halko HM, Nolen EA, Powell BJ, Dorsey CN, Mettert KD, et al. Pragmatic measures for implementation research: development of the Psychometric and Pragmatic Evidence Rating Scale. Transl Behav Med. 2021;11(1):11–20.
44. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res. 2010;19(4):539–49.
45. Terwee CB, Prinsen CA, Chiarotto A, Westerman M, Patrick DL, Alonso J, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. Qual Life Res. 2018;27(5):1159–70.
46. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Ann Intern Med*. 2018;169(7):467–73.

ARTICLE IN PRESS

47. Elsman EBM, Mokkink LB, Terwee CB, Beaton D, Gagnier JJ, Tricco AC, et al. Guideline for reporting systematic reviews of outcome measurement instruments (OMIs): PRISMA-COSMIN for OMIs 2024. J Clin Epidemiol. 2024;173:111422.
48. Squires J, Hutchinson A, Graham I, Santos W, team I. Review of Tools that Could be Used to Measure Context in Healthcare: OSF; 2022 [Available from: https://osf.io/b2z8k.
49. McGowan J, Sampson M, Lefebvre C. An Evidence Based Checklist for the Peer Review of Electronic Search Strategies (PRESS EBC). Evid Based Libr Inf Pract. 2010;5.1:149–54.
50. McMaster University Health Information Research Unit. Search Filters For Medline In Ovid Syntax And The Pubmed Translation: McMaster University 2022 [Available from: https://hiruweb.mcmaster.ca/hkr/hedges/medline/.
51. Bramer WM, de Jonge GB, Rethlefsen ML, Mast F, Kleijnen J. A systematic approach to searching: an efficient and complete method to develop literature searches. J Med Libr Assoc. 2018;106(4):531–41.
52. Clarivate. EndNote 2023 [Available from: https://endnote.com/?language=en.
53. DistillerSR. DistillerSR: DistillerSR; 2023 [Available from: https://www.distillersr.com/products/distillersr-systematic-review-software.
54. Terwee CB, Jansma EP, Riphagen II, de Vet HC. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. Qual Life Res. 2009;18(8):1115–23.
55. Santos WJ, Hutchinson AM, Rader T, Graham ID, Watkins V, Candido LKd, et al. Insights from using an outcomes measurement properties search filter and conducting citation searches to locate psychometric articles of tools used to measure context attributes. BMC Res Notes. 2023;16(1):34.
56. Terwee CB, Prinsen C, Chiarotto A, De Vet H, Bouter LM, Alonso J, et al. COSMIN methodology for assessing the content validity of PROMs–user manual 2018 [Available from: https://www.cosmin.nl/wp-

ARTICLE IN PRESS

content/uploads/COSMIN-methodology-for-content-validity-user-manual-v1.pdf.
57. Mokkink LB, Prinsen C, Patrick DL, Alonso J, Bouter LM, De Vet H, et al. COSMIN methodology for systematic reviews of patient-reported outcome measures (PROMs). User manual. 2018;78(1):6–3.
58. Hatcher M. Survey of acute care hospitals in the United States relative to technology usage and technology transfer. Journal of Medical Systems. 1997;21:323–36.
59. Webster G, Swanson D. Final report on the patient satisfaction questionnaire project. Philadelphia: American Board of Internal Medicine. 1989.
60. Stump TE, Dexter PR, Tierney WM, Wolinsky FD. Measuring Patient Satisfaction with Physicians among Older and Diseased Adults in a Primary Care Municipal Outpatient Setting An Examination of Three Instruments. Medical Care. 1995;33(9):973–2.
61. Stogdill RM. Handbook of leadership: A survey of theory and research: Free Press; 1974.
62. Anderson N, West MA. The Team Climate Inventory: Development of the TCI and its applications in teambuilding for innovativeness. Eur J Work Organ. 1996;5(1):53–66.
63. Anderson NR, West MA. Measuring climate for work group innovation: development and validation of the team climate inventory. J Organ Behav. 1998;19(3):235–58.
64. Stewart AL, Nápoles-Springer A, Pérez-Stable EJ. Interpersonal processes of care in diverse populations. Milbank Q. 1999;77(3):305–39, 274.
65. Stewart AL, Nápoles-Springer AM, Gregorich SE, Santoyo-Olsson J. Interpersonal processes of care survey: patient-reported measures for diverse groups. Health Serv Res. 2007;42(3 Pt 1):1235–56.
66. Körner M, Wirtz MA. Development and psychometric properties of a scale for measuring internal participation from a patient and health care professional perspective. BMC Health Serv Res. 2013;13:374.
67. Nahm ES, Resnick B, Mills ME. Development and pilot-testing of the perceived health Web Site usability questionnaire (PHWSUQ) for older adults. Stud Health Technol Inform. 2006;122:38–43.

ARTICLE IN PRESS

ARTICLE IN PRESS

68. Murphy CA, Maynard M, Morgan G. Pretest and post-test attitudes of nursing personnel toward a patient care information system. Computers in Nursing. 1994;12(5):239–44.
69. Weiss SJ, Davis HP. Validity and reliability of the Collaborative Practice Scales. Nurs Res. 1985;34(5):299–305.
70. Kenaszchuk C, Reeves S, Nicholas D, Zwarenstein M. Validity and reliability of a multiple-group measurement scale for interprofessional collaboration. BMC Health Serv Res. 2010;10:83.
71. Scholl I, Kriston L, Dirmaier J, Buchholz A, Härter M. Development and psychometric properties of the Shared Decision Making Questionnaire--physician version (SDM-Q-Doc). Patient Educ Couns. 2012;88(2):284–90.
72. Howell JM, Shea CM, Higgins CA. Champions of product innovations: Defining, developing, and validating a measure of champion behavior. Journal of business venturing. 2005;20(5):641–61.
73. Oreg S. Resistance to change: developing an individual differences measure. J Appl Psychol. 2003;88(4):680–93.
74. Greco M, Cavanagh M, Brownlea A, McGovern J. The Doctor's Interpersonal Skills Questionnaire (DISQ): a validated instrument for use in GP training. Education for General Practice. 1999;10:256–64.
75. Ehrhart M, Naumann S. Leadership and procedural justice climate citizenship behavior as antecedents of unit-level organizational. Personnel Psychology. 2004;57(1):61–94.
76. Carless SA, Wearing AJ, Mann L. A short measure of transformational leadership. J Bus Psychol. 2000;14:389–405.
77. Aarons GA, Ehrhart MG, Farahnak LR. The Implementation Leadership Scale (ILS): development of a brief measure of unit level implementation leadership. Implement Sci. 2014;9(1):45.
78. Barbuto Jr JE, Wheeler DW. Scale development and construct clarification of servant leadership. Group Organ Manag. 2006;31(3):300–26.
79. van Dierendonck D, Nuijten I. The Servant Leadership Survey: Development and Validation of a Multidimensional Measure. J Bus Psychol. 2010;26(3):249–67.

ARTICLE IN PRESS

ARTICLE IN PRESS

80. Berk RA, Berg J, Mortimer R, Walton-Moss B, Yeo TP. Measuring the effectiveness of faculty mentoring relationships. Acad Med. 2005;80(1):66–71.
81. Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research, 1988. Advances in Psychology. 1988.
82. Costa AC, Anderson N. Measuring trust in teams: Development and validation of a multifaceted measure of formative and reflective indicators of team trust. Eur J Work Organ. 2011;20(1):119–54.
83. Jones J, Barry MM. Developing a scale to measure trust in health promotion partnerships. Health Promot Int. 2011;26(4):484–91.
84. Singer S, Meterko M, Baker L, Gaba D, Falwell A, Rosen A. Workforce perceptions of hospital safety culture: development and validation of the patient safety climate in healthcare organizations survey. Health Serv Res. 2007;42(5):1999–2021.
85. Schutz AL, Counte MA, Meurer S. Development of a patient safety culture measurement tool for ambulatory health care settings: analysis of content validity. Health Care Manag Sci. 2007;10(2):139–49.
86. Shea CM, Jacobs SR, Esserman DA, Bruce K, Weiner BJ. Organizational readiness for implementing change: a psychometric assessment of a new measure. Implement Sci. 2014;9:7.
87. Frank C, Asp M, Fridlund B, Baigi A. Questionnaire for patient participation in emergency departments: development and psychometric testing. J Adv Nurs. 2011;67(3):643–51.
88. Risser NL. Development of an instrument to measure patient satisfaction with nurses and nursing care in primary care settings. Nurs Res. 1975;24(1):45–52.
89. Welbourne TM, Johnson DE, Erez A. The role-based performance scale: Validity analysis of a theory-based measure. Acad Manage J. 1998;41(5):540–55.
90. Simon D, Schorr G, Wirtz M, Vodermaier A, Caspari C, Neuner B, et al. Development and first validation of the shared decision-making questionnaire (SDM-Q). Patient Educ Couns. 2006;63(3):319–27.

ARTICLE IN PRESS

91. Glasgow RE, Wagner EH, Schaefer J, Mahoney LD, Reid RJ, Greene SM. Development and validation of the Patient Assessment of Chronic Illness Care (PACIC). Medical Care. 2005;43(5):436–44.
92. O'Cathain A, Knowles E, Nicholl J. Measuring patients' experiences and views of the emergency and urgent care system: psychometric testing of the urgent care system questionnaire. BMJ Qual Saf. 2011;20(2):134–40.
93. Fagerström L, Rainio AK. Professional assessment of optimal nursing care intensity level: a new method of assessing personnel resources for nursing care. J Clin Nurs. 1999;8(4):369–79.
94. King G, Servais M, Kertoy M, Specht J, Currie M, Rosenbaum P, et al. A measure of community members' perceptions of the impacts of research partnerships in health and social services. Eval Program Plann. 2009;32(3):289–99.
95. Elwyn G, Barr PJ, Grande SW, Thompson R, Walsh T, Ozanne EM. Developing CollaboRATE: a fast and frugal patient-reported measure of shared decision making in clinical encounters. Patient Educ Couns. 2013;93(1):102–7.
96. Batorowicz B, Shepherd TA. Measuring the quality of transdisciplinary teams. J Interprof Care 2008;22(6):612–20.
97. Coleman EA, Smith JD, Frank JC, Eilertsen TB, Thiare JN, Kramer AM. Development and testing of a measure designed to assess the quality of care transitions. Int J Integr Care. 2002;2:e02.
98. Ouwens M, Hulscher M, Akkermans R, Hermens R, Grol R, Wollersheim H. The Team Climate Inventory: application in hospital teams and methodological considerations. Qual Saf Health Care. 2008;17(4):275–80.
99. Kivimaki M, Elovainio M. A short version of the Team Climate Inventory: Development and psychometric properties. J Occup Psychol. 1999;72(2):241–6.
100. Kivimaki M, Kuk G, Elovainio M, Thomson L, Kalliomaki-Levanto T, Heikkila A. The Team Climate Inventory (TCI)- four or five factors? Testing the structure of TCI in smaples of low and high complexity jobs. J Occup Psychol. 1997;70:375–89.

ARTICLE IN PRESS

ARTICLE IN PRESS

101. Beaulieu MD, Dragieva N, Del Grande C, Dawson J, Haggerty JL, Barnsley J, et al. The team climate inventory as a measure of primary care teams' processes: validation of the French version. Healthc Policy. 2014;9(3):40–54.
102. Strating MM, Nieboer AP. Psychometric test of the Team Climate Inventory-short version investigated in Dutch quality improvement teams. BMC Health Serv Res. 2009;9(1):1–7.
103. Kivimaki M, Elovainio M. A short version of the Team Climate Inventory: Development and psychometric properties. J Occup Organ Psychol. 1999;72(2):241–6.
104. Silva MC, Peduzzi M, Sangaleti CT, Silva D, Agreli HF, West MA, et al. Cross-cultural adaptation and validation of the teamwork climate scale. Rev Saude Publica. 2016;50:52.
105. Son H, Nahm ES, Zhu S, Galik E, Seidl KL, Van de Castle B, et al. Testing a Model of Patient Portal Use in Adult Patients. J Nurs Scholarsh. 2021;53(2):143–53.
106. Goto Y, Yamaguchi Y, Onishi J, Arai H, Härter M, Scholl I, et al. Adapting the patient and physician versions of the 9-item shared decision making questionnaire for other healthcare providers in Japan. BMC Med Inform Decis Mak. 2021;21(1):314.
107. Howell JM, Shea CM, Higgins CA. Champions of product innovations: Defining, developing, and validating a measure of champion behavior. J Bus Ventur. 2005;20(5):641–61.
108. Lindig A, Hahlweg P, Christalle E, Scholl I. Translation and psychometric evaluation of the German version of the Organisational Readiness for Implementing Change measure (ORIC): a cross-sectional study. BMJ Open. 2020;10(6):e034380.
109. Storkholm MH, Mazzocato P, Tessma MK, Savage C. Assessing the reliability and validity of the Danish version of Organizational Readiness for Implementing Change (ORIC). Implement Sci. 2018;13(1):78.
110. Bomfim RA, Braff EC, Frazão P. Cross-cultural adaptation and psychometric properties of the Brazilian-Portuguese version of the Organizational Readiness for Implementing Change questionnaire. Rev Bras Epidemiol. 2020;23:e200100.

ARTICLE IN PRESS

ARTICLE IN PRESS

111. Ruest M, Léonard G, Thomas A, Desrosiers J, Guay M. French cross-cultural adaptation of the Organizational Readiness for Implementing Change (ORIC). BMC Health Serv Res. 2019;19(1):535.
112. Leslie HH, West R, Twine R, Masilela N, Steward WT, Kahn K, et al. Measuring Organizational Readiness for Implementing Change in Primary Care Facilities in Rural Bushbuckridge, South Africa. Int J Health Policy Manag. 2022;11(7):912–8.
113. Oreg S, Bayazit M, Vakola M, Arciniega L, Armenakis A, Barkauskiene R, et al. Dispositional resistance to change: Measurement equivalence and the link to personal values across 17 nations. J Appl Psychol. 2008;93(4):935.
114. Nápoles-Springer AM, Santoyo-Olsson J, O'Brien H, Stewart AL. Using cognitive interviews to develop surveys in diverse populations. Med Care. 2006;44(11 Suppl 3):S21–30.
115. Peimani M, Nasli-Esfahani E, Stewart AL, Shakibazadeh E, Yekaninejad MS, Garmaroudi G. Patient-physician communication in the context of diabetes care: Adaptation and assessment of psychometric properties of a Persian version of the interpersonal process of care survey (P-IPC). Diabetes Res Clin Pract. 2020;162:108109.
116. Shiozawa T, Yamaguchi S, Matsunaga A, Sawada U, Fujii C. Development of the Interpersonal Processes of Care Survey-Japanese version. Neuropsychopharmacol Rep. 2020;40(1):107–12.
117. Beaulieu MD, Haggerty JL, Beaulieu C, Bouharaoui F, Lévesque JF, Pineault R, et al. Interpersonal communication from the patient perspective: comparison of primary healthcare evaluation instruments. Healthc Policy. 2011;7(Spec Issue):108–23.
118. Haggerty JL, Beaulieu C, Lawson B, Santor DA, Fournier M, Burge F. What Patients Tell Us about Primary Healthcare Evaluation Instruments: Response Formats, Bad Questions and Missing Pieces. Healthc Policy. 2011;7(Spec Issue):66–78.
119. Haggerty JL, Bouharaoui F, Santor DA. Differential item functioning in primary healthcare evaluation instruments by french/english version, educational level and urban/rural location. Healthc Policy. 2011;7(Spec Issue):47–65.

ARTICLE IN PRESS

120. Haggerty JL, Burge F, Beaulieu MD, Pineault R, Beaulieu C, Lévesque JF, et al. Validation of instruments to evaluate primary healthcare from the patient perspective: overview of the method. Healthc Policy. 2011;7(Spec Issue):31–46.
121. Alvarado-Villa GE, Moncayo-Rizzo JD, Gallardo-Rumbea JA. Spanish validation endorsement of SDM-Q-9, a new approach. BMC Public Health. 2019;19(1):106.
122. Alvarez K, Wang Y, Alegria M, Ault-Brutus A, Ramanayake N, Yeh YH, et al. Psychometrics of shared decision making and communication as patient centered measures for two language groups. Psychol Assess. 2016;28(9):1074–86.
123. Baicus C, Balanescu P, Gurghean A, Badea CG, Padureanu V, Rezus C, et al. Romanian version of SDM-Q-9 validation in Internal Medicine and Cardiology setting: a multicentric cross-sectional study. Rom J Intern Med. 2019;57(2):195–200.
124. Ballesteros J, Moral E, Brieva L, Ruiz-Beato E, Prefasi D, Maurino J. Psychometric properties of the SDM-Q-9 questionnaire for shared decision-making in multiple sclerosis: item response theory modelling and confirmatory factor analysis. Health Qual Life Outcomes. 2017;15(1):79.
125. Calderon C, Jiménez-Fonseca P, Ferrando PJ, Jara C, Lorenzo-Seva U, Beato C, et al. Psychometric properties of the Shared Decision-Making Questionnaire (SDM-Q-9) in oncology practice. Int J Clin Health Psychol. 2018;18(2):143–51.
126. Hurley EA, Bradley-Ewing A, Bickford C, Lee BR, Myers AL, Newland JG, et al. Measuring shared decision-making in the pediatric outpatient setting: Psychometric performance of the SDM-Q-9 and CollaboRATE among English and Spanish speaking parents in the US Midwest. Patient Educ Couns. 2019;102(4):742–8.
127. Kasper J, Heesen C, Köpke S, Fulcher G, Geiger F. Patients' and observers' perceptions of involvement differ. Validation study on inter-relating measures for shared decision making. PLoS ONE. 2011;6(10):e26255.
128. Nejati B, Lin CC, Imani V, Browall M, Lin CY, Broström A, et al. Validating patient and physician versions of the shared decision making

ARTICLE IN PRESS

questionnaire in oncology setting. Health Promot Perspect. 2019;9(2):105–14.
129. Rodenburg-Vandenbussche S, Pieterse AH, Kroonenberg PM, Scholl I, van der Weijden T, Luyten GP, et al. Dutch Translation and Psychometric Testing of the 9-Item Shared Decision Making Questionnaire (SDM-Q-9) and Shared Decision Making Questionnaire-Physician Version (SDM-Q-Doc) in Primary and Secondary Care. PLoS ONE. 2015;10(7):e0132158.
130. Scholl I, Kriston L, Dirmaier J, Härter M. Comparing the nine-item Shared Decision-Making Questionnaire to the OPTION Scale - an attempt to establish convergent validity. Health Expect. 2015;18(1):137–50.
131. Schulz AA, Wirtz MA. Assessment of the quality of woman-centred midwifery care from the mothers' perspective: A structural analysis of cross-sectional survey data. Z Evid Fortbild Qual Gesundhwes. 2021;166:8–17.
132. Shunnmugam B, Ng CJ, Aishah Mohd Taib N, Chinna K. Validation of the Malay, English, and Chinese Translations of the 9-Item Shared Decision Making Questionnaire (SDM-Q-9) in Breast Cancer Patients Making Treatment Decisions. Asia Pac J Public Health. 2021;33(8):914–22.
133. Wu TY, Chen CT, Huang YJ, Hou WH, Wang JD, Hsieh CL. Rasch Analysis of the 9-Item Shared Decision Making Questionnaire in Women With Breast Cancer. Cancer Nurs. 2019;42(3):E34–e42.
134. Zisman-Ilani Y, Roe D, Scholl I, Härter M, Karnieli-Miller O. Shared Decision Making During Active Psychiatric Hospitalization: Assessment and Psychometric Properties. Health Commun. 2017;32(1):126–30.
135. Alzubaidi H, Hussein A, Mc Namara K, Scholl I. Psychometric properties of the Arabic version of the 9-item Shared Decision-Making Questionnaire: the entire process from translation to validation. BMJ Open. 2019;9(4):e026672.
136. De las Cuevas C, Perestelo-Perez L, Rivero-Santana A, Cebolla-Martí A, Scholl I, Härter M. Validation of the Spanish version of the 9-

ARTICLE IN PRESS

item Shared Decision-Making Questionnaire. Health Expect. 2015;18(6):2143–53.
137. Goto Y, Miura H, Son D, Arai H, Kriston L, Scholl I, et al. Psychometric Evaluation of the Japanese 9-Item Shared Decision-Making Questionnaire and Its Association with Decision Conflict and Patient Factors in Japanese Primary Care. Jma J. 2020;3(3):208–15.
138. Kriston L, Härter M, Scholl I. A latent variable framework for modeling dyadic measures in research on shared decision-making. Z Evid Fortbild Qual Gesundhwes. 2012;106(4):253–63.
139. Kriston L, Scholl I, Hölzel L, Simon D, Loh A, Härter M. The 9-item Shared Decision Making Questionnaire (SDM-Q-9). Development and psychometric properties in a primary care sample. Patient Educ Couns. 2010;80(1):94–9.
140. Rencz F, Tamási B, Brodszky V, Gulácsi L, Weszl M, Péntek M. Validity and reliability of the 9-item Shared Decision Making Questionnaire (SDM-Q-9) in a national survey in Hungary. Eur J Health Econ. 2019;20(Suppl 1):43–55.
141. Aragones A, Schaefer EW, Stevens D, Gourevitch MN, Glasgow RE, Shah NR. Validation of the Spanish translation of the Patient Assessment of Chronic Illness Care (PACIC) survey. Prev Chronic Dis. 2008;5(4):A113.
142. Aung E, Ostini R, Dower J, Donald M, Coll JR, Williams GM, et al. Patient Assessment of Chronic Illness Care (PACIC) in Type 2 Diabetes: A Longitudinal Study. Eval Health Prof. 2016;39(2):185–203.
143. Cramm JM, Nieboer AP. Factorial validation of the Patient Assessment of Chronic Illness Care (PACIC) and PACIC short version (PACIC-S) among cardiovascular disease patients in the Netherlands. Health Qual Life Outcomes. 2012;10:104.
144. Gugiu PC, Coryn C, Clark R, Kuehn A. Development and evaluation of the short version of the Patient Assessment of Chronic Illness Care instrument. Chronic Illn. 2009;5(4):268–76.
145. Iglesias K, Burnand B, Peytremann-Bridevaux I. PACIC Instrument: disentangling dimensions using published validation models. Int J Qual Health Care. 2014;26(3):250–60.

ARTICLE IN PRESS

146. Kim SH, Lee BG, Choe YH. Psychometric evaluation of patient assessment of chronic illness care among Korean cancer survivors. PLoS ONE. 2021;16(8):e0256119.
147. Koley M, Saha S, Ghosh S, Nag G, Kundu M, Mondal R, et al. Patient-Assessed Chronic Illness Care (PACIC) scenario in an Indian homeopathic hospital. J Tradit Complement Med. 2016;6(1):72–7.
148. Krucien N, Le Vaillant M, Pelletier-Fleury N. Adaptation and validation of the patient assessment of chronic illness care in the French context. BMC Health Serv Res. 2014;14:269.
149. Maindal HT, Sokolowski I, Vedsted P. Adaptation, data quality and confirmatory factor analysis of the Danish version of the PACIC questionnaire. Eur J Public Health. 2010;22(1):31–6.
150. Simonsen N, Koponen AM, Suominen S. Patients' assessment of chronic illness care: a validation study among patients with type 2 diabetes in Finland. BMC Health Serv Res. 2018;18(1):412.
151. Cramm JM, Nieboer AP. Development and validation of the older patient assessment of chronic illness care (O-PACIC) scale after hospitalization. Soc Indic Res. 2014;116:959–69.
152. Abdul-Razak S, Ramli AS, Badlishah-Sham SF, Haniff J. Validity and reliability of the patient assessment on chronic illness care (PACIC) questionnaire: the Malay version. BMC Fam Pract. 2018;19(1):119.
153. Azam AF, Lai P, Abdullah A, Haidi Hanafi NS. Validity and reliability of the Patient Assessment on Chronic Illness Care (PACIC) questionnaire among patients with type 2 diabetes mellitus in Malaysia: English version. Malays Fam Physician. 2020;15(2):10–8.
154. Fan J, McCoy RG, Ziegenfuss JY, Smith SA, Borah BJ, Deming JR, et al. Evaluating the structure of the Patient Assessment of Chronic Illness Care (PACIC) survey from the patient's perspective. Ann Behav Med. 2015;49(1):104–11.
155. Gensichen J, Serras A, Paulitsch MA, Rosemann T, König J, Gerlach FM, et al. The Patient Assessment of Chronic Illness Care questionnaire: evaluation in patients with mental disorders in primary care. Community Ment Health J. 2011;47(4):447–53.
156. Gibbons CJ, Small N, Rick J, Burt J, Hann M, Bower P. The Patient Assessment of Chronic Illness Care produces measurements

ARTICLE IN PRESS

along a single dimension: results from a Mokken analysis. Health Qual Life Outcomes. 2017;15(1):61.
157. Noël PH, Jones S, Parchman ML. Patient Experience in an Era of Primary Care Transformation: Revisiting the PACIC. Eur J Pers Cent Healthc. 2016;4(3):528–40.
158. Rick J, Rowe K, Hann M, Sibbald B, Reeves D, Roland M, et al. Psychometric properties of the Patient Assessment Of Chronic Illness Care measure: acceptability, reliability and validity in United Kingdom patients with long-term conditions. BMC Health Serv Res. 2012;12:293.
159. Taggart J, Chan B, Jayasinghe UW, Christl B, Proudfoot J, Crookes P, et al. Patients Assessment of Chronic Illness Care (PACIC) in two Australian studies: structure and utility. J Eval Clin Pract. 2011;17(2):215–21.
160. Tusek-Bunc K, Petek-Ster M, Ster B, Petek D, Kersnik J. Validation of the Slovenian version of patient assessment of chronic illness care (PACIC) in patients with coronary heart disease. Coll Antropol. 2014;38(2):437–44.
161. Wensing M, van Lieshout J, Jung HP, Hermsen J, Rosemann T. The Patients Assessment Chronic Illness Care (PACIC) questionnaire in The Netherlands: a validation study in rural general practice. BMC Health Serv Res. 2008;8:182.
162. Ali S, Chessex C, Bassett-Gunter R, Grace SL. Patient satisfaction with cardiac rehabilitation: association with utilization, functional capacity, and heart-health behaviors. Patient Prefer Adherence. 2017;11:821–30.
163. Lambert S, McCusker J, Belzile E, Yaffe M, Ihejirika C, Richardson J, et al. Using confirmatory factor analysis and Rasch analysis to examine the dimensionality of The Patient Assessment of Care for Chronic Illness Care (PACIC). Qual Life Res. 2021;30(5):1503–12.
164. Iglesias K, De Geest S, Berben L, Dobbels F, Denhaerynk K, Russell LC, et al. Validation of the patient assessment of chronic illness care (PACIC) short form scale in heart transplant recipients: the international cross-sectional bright study. BMC Health Serv Res. 2020;20(1):160.

ARTICLE IN PRESS

165. Goetz K, Freund T, Gensichen J, Miksch A, Szecsenyi J, Steinhaeuser J. Adaptation and psychometric properties of the PACIC short form. Am J Manag Care. 2012;18(2):e55–60.
166. Vennedey V, Dust G, Schippel N, Shukri A, Strupp J, Rietz C, et al. Patient-centered care during the last year of life: adaptation and validation of the German PACIC short form for bereaved persons as proxies (PACIC-S9-proxy). BMC Palliat Care. 2020;19(1):177.
167. Alharbi N, Alsubki N, Alotabi F, Alotabi M, Alhrabi N, de Lusgnian S, et al. Translation into Arabic and validation of the Patient Assessment of Care for Chronic Conditions questionnaire for diabetes. East Mediterr Health J. 2021;27(2):142–50.
168. Malliarou M, Bakola E, Nikolentzos A, Sarafis P. Reliability and validity of the Greek translation of the patient assessment of chronic illness care + (PACIC-PLUS GR) survey. BMC Fam Pract. 2020;21(1):122.
169. Zeugfang D, Wisetborisut A, Angkurawaranon C, Aramrattana A, Wensing M, Szecsenyi J, et al. Translation and validation of the PACIC+ questionnaire: the Thai version. BMC Fam Pract. 2018;19(1):123.
170. Drewes HW, de Jong-van Til JT, Struijs JN, Baan CA, Tekle FB, Meijboom BR, et al. Measuring chronic care management experience of patients with diabetes: PACIC and PACIC+ validation. Int J Integr Care. 2012;12:e194.
171. Rosemann T, Laux G, Droesemeyer S, Gensichen J, Szecsenyi J. Evaluation of a culturally adapted German version of the Patient Assessment of Chronic Illness Care (PACIC 5A) questionnaire in a sample of osteoarthritis patients. J Eval Clin Pract. 2007;13(5):806–13.
172. Schwenke M, Welzel FD, Luck-Sikorski C, Pabst A, Kersting A, Blüher M, et al. Psychometric properties of the Patient Assessment of Chronic Illness Care measure (PACIC-5A) among patients with obesity. BMC Health Serv Res. 2019;19(1):61.
173. Acosta AM, Lima M, Marques GQ, Levandovski PF, Weber LAF. Brazilian version of the Care Transitions Measure: translation and validation. Int Nurs Rev. 2017;64(3):379–87.

ARTICLE IN PRESS

ARTICLE IN PRESS

174. Anatchkova MD, Barysauskas CM, Kinney RL, Kiefe CI, Ash AS, Lombardini L, et al. Psychometric evaluation of the Care Transition Measure in TRACE-CORE: do we need a better measure? J Am Heart Assoc. 2014;3(3):e001053.
175. Flink M, Tessma M, Cvancarova Småstuen M, Lindblad M, Coleman EA, Ekstedt M. Measuring care transitions in Sweden: validation of the care transitions measure. Int J Qual Health Care. 2018;30(4):291–7.
176. Parry C, Mahoney E, Chalmers SA, Coleman EA. Assessing the quality of transitional care: further applications of the care transitions measure. Med Care. 2008;46:317–22.
177. Bakshi AB, Wee SL, Tay C, Wong LM, Leong IY, Merchant RA, et al. Validation of the care transition measure in multi-ethnic South-East Asia in Singapore. BMC Health Serv Res. 2012;12:256.
178. Cao X, Chen L, Diao Y, Tian L, Liu W, Jiang X. Validity and reliability of the Chinese version of the care transition measure. PLoS ONE. 2015;10(5):e0127403.
179. Shadmi E, Zisberg A, Coleman EA. Translation and validation of the Care Transition Measure into Hebrew and Arabic. Int J Qual Health Care. 2009;21(2):97–102.
180. Coleman EA, Mahoney E, Parry C. Assessing the quality of preparation for posthospital care from the patient's perspective: the care transitions measure. MedCare. 2005;43:246–55.
181. Hwang JI, Chung JH, Kim HK. Psychometric properties of transitional care instruments and their relationships with health literacy: Brief PREPARED and Care Transitions Measure. Int J Qual Health Care. 2019;31(10):774–80.
182. Yoshimura M, Sato M, Sumi N. Validity and reliability of the Japanese version of the Care Transitions Measure. Int J Health Plann Manage. 2017;33(2):380–90.
183. McLeod J, Stolee P, Walker J, Heckman G. Measuring care transition quality for older patients with musculoskeletal disorders. Musculoskelet Care. 2013;12(1):13–21.
184. Aarons GA, Ehrhart MG, Torres EM, Finn NK, Roesch SC. Validation of the Implementation Leadership Scale (ILS) in Substance

ARTICLE IN PRESS

use Disorder Treatment Organizations. J Subst Abuse Treat. 2016;68:31–5.
185. Hu J, Gifford W, Ruan H, Harrison D, Li Q, Ehrhart MG, et al. Translation and linguistic validation of the implementation leadership scale in Chinese nursing context. J Nurs Manag. 2019;27(5):1030–8.
186. Hu J, Gifford W, Ruan H, Harrison D, Li Q, Ehrhart MG, et al. Validating the Implementation Leadership Scale in Chinese nursing context: A cross-sectional study. Nurs Open. 2021;8(6):3420–9.
187. Mandrou E, Tsounis A, Sarafis P. Validity and reliability of the Greek version of Implementation Leadership Scale (ILS). BMC Psychol. 2020;8(1):49.
188. Shuman CJ, Ehrhart MG, Torres EM, Veliz P, Kath LM, VanAntwerp K, et al. EBP Implementation Leadership of Frontline Nurse Managers: Validation of the Implementation Leadership Scale in Acute Care. Worldviews Evid Based Nurs. 2020;17(1):82–91.
189. Torres EM, Ehrhart MG, Beidas RS, Farahnak LR, Finn NK, Aarons GA. Validation of the Implementation Leadership Scale (ILS) with Supervisors' Self-Ratings. Community Ment Health J. 2018;54(1):49–53.
190. Kobayashi Y, Watanabe K, Otsuka Y, Eguchi H, Kawakami N, Imamura K, et al. Servant Leadership in Japan: A Validation Study of the Japanese Version of the Servant Leadership Survey (SLS-J). Front Psychol. 2020;11:1711.
191. Rodríguez-Carvajal R, de Rivas S, Herrero M, Moreno-Jiménez B, van Dierendonck D. Leading people positively: cross-cultural validation of the Servant Leadership Survey (SLS). Span J Psychol. 2014;17:E63.
192. Hernandez R, Roll SC, Jin H, Schneider S, Pyatak EA. Validation of the National Aeronautics and Space Administration Task Load Index (NASA-TLX) adapted for the whole day repeated measures context. Ergonomics. 2021;65(7):960–75.
193. Tubbs-Cooley HL, Mara CA, Carle AC, Gurses AP. The NASA Task Load Index as a measure of overall workload among neonatal, paediatric and adult intensive care nurses. Intensive Crit Care Nurs. 2018;46:64–9.

ARTICLE IN PRESS

ARTICLE IN PRESS

194. Frilund M, Fagerström L. Validity and reliability testing of the Oulu patient classification: instrument within primary health care for the older people. Int J Older People Nurs. 2009;4(4):280–7.
195. Rauhala A, Fagerström L. Determining optimal nursing intensity: the RAFAELA method. J Adv Nurs. 2004;45(4):351–9.
196. Junttila JK, Koivu A, Fagerström L, Haatainen K, Nykänen P. Hospital mortality and optimality of nursing workload: A study on the predictive validity of the RAFAELA Nursing Intensity and Staffing system. Int J Nurs Stud. 2016;60:46–53.
197. Fagerström L, Rainio AK, Rauhala A, Nojonen K. Validation of a new method for patient classification, the Oulu Patient Classification. J Adv Nurs. 2000;31(2):481–90.
198. Andersen MH, Lønning K, Bjørnelv GM, Fagerström L. Nursing intensity and costs of nurse staffing demonstrated by the RAFAELA system: liver vs. kidney transplant recipients. J Nurs Manag. 2016;24(6):798–805.
199. Flo J, Landmark B, Hatlevik OE, Fagerström L. Using a new interrater reliability method to test the modified Oulu Patient Classification instrument in home health care. Nurs Open. 2018;5(2):167–75.
200. Borsboom D, Mellenbergh GJ, Van Heerden J. The concept of validity. Psychol Rev. 2004;111(4):1061.
201. De Vet HC, Terwee CB, Mokkink LB, Knol DL. Measurement in medicine: a practical guide: Cambridge University Press; 2011.
202. Wang Y, Yin M, Zhu S, Chen X, Zhou H, Qian W. Patient-reported outcome measures used in patients undergoing total knee arthroplasty: a COSMIN systematic review. Bone Jt Res. 2021;10(3):203–17.
203. Rezai M, Kolne K, Bui S, Lindsay S. Measures of workplace inclusion: a systematic review using the COSMIN methodology. J Occup Rehabil. 2020;30(3):420–54.
204. Hoglund LT, Scalzitti DA, Bolgla LA, Jayaseelan DJ, Wainwright SF. Patient-Reported Outcome Measures for Adults and Adolescents with Patellofemoral Pain: A Systematic Review of Content Validity and

ARTICLE IN PRESS

Feasibility Using the COSMIN Methodology. J Orthop Sports Phys Ther. 2023;53(1):23–39.
205. Foley T, Vale L. A framework for understanding, designing, developing and evaluating learning health systems. Learn Health Syst. 2023;7(1):e10315.
206. Strating MM, Nieboer AP. Psychometric test of the Team Climate Inventory-short version investigated in Dutch quality improvement teams. BMC Health Serv Res. 2009;9:126.
207. Hatcher M. Survey of acute care hospitals in the United States relative to technology usage and technology transfer. *J Med Syst*. 1997;21(5):323–36.
208. Rettke H, Lehmann AI, Brauchli R, Bauer GF, Petry H, Spirig R. Capturing interprofessional collaboration between physicians and nurses in an acute care setting. A validation study of the revised German version of the Collaborative Practice Scales. J Interprof Care. 2020;34(2):211–7.
209. Vittadello F, Mischo-Kelling M, Wieser H, Cavada L, Lochner L, Naletto C, et al. A multiple-group measurement scale for interprofessional collaboration: Adaptation and validation into Italian and German languages. J Interprof Care. 2018;32(3):266–73.
210. Kenaszchuk C, Wilkins K, Reeves S, Zwarenstein M, Russell A. Nurse-physician relations and quality of nursing care: findings from a national survey of nurses. Canadian Journal of Nursing Research. 2010;42(2):120–36.
211. Calderon C, Ferrando PJ, Carmona-Bayonas A, Lorenzo-Seva U, Jara C, Beato C, et al. Validation of SDM-Q-Doc Questionnaire to measure shared decision-making physician's perspective in oncology practice. Clin Transl Oncol. 2017;19(11):1312–9.
212. Goto Y, Miura H, Son D, Scholl I, Kriston L, Härter M, et al. Association between physicians' and patients' perspectives of shared decision making in primary care settings in Japan: The impact of environmental factors. PLoS ONE. 2021;16(2):e0246518.
213. Oreg S, Bayazit M, Vakola M, Arciniega L, Armenakis A, Barkauskiene R, et al. Dispositional resistance to change: measurement

ARTICLE IN PRESS

equivalence and the link to personal values across 17 nations. J Appl Psychol. 2008;93(4):935–44.
214. Stump TE, Dexter PR, Tierney WM, Wolinsky FD. Measuring patient satisfaction with physicians among older and diseased adults in a primary care municipal outpatient setting. An examination of three instruments. Med Care. 1995;33(9):958–72.
215. Yoshimura M, Sato M, Sumi N. Validity and reliability of the Japanese version of the Care Transitions Measure. Int J Health Plann Manage. 2018;33(2):380–90.
216. Ruiz Yanzi MV, Barani MS, Franco JVA, Vazquez Peña FR, Terrasa SA, Kopitowski KS. Translation, transcultural adaptation, and validation of two questionnaires on shared decision making. Health Expect. 2019;22(2):193–200.
217. Broström A, Pakpour AH, Nilsen P, Hedberg B, Ulander M. Validation of CollaboRATE and SURE - two short questionnaires to measure shared decision making during CPAP initiation. J Sleep Res. 2019;28(5):e12808.
218. Ubbink DT, van Asbeck EV, Aarts JWM, Stubenrouch FE, Geerts PAF, Atsma F, et al. Comparison of the CollaboRATE and SDM-Q-9 questionnaires to appreciate the patient-reported level of shared decision-making. Patient Educ Couns. 2022;105(7):2475–9.
219. Ousseine YM, Durand MA, Bouhnik AD, Smith A, Mancini J. Multiple health literacy dimensions are associated with physicians' efforts to achieve shared decision-making. Patient Educ Couns. 2019;102(11):1949–56.
220. Brodney S, Fowler FJ, Jr., Barry MJ, Chang Y, Sepucha K. Comparison of Three Measures of Shared Decision Making: SDM Process_4, CollaboRATE, and SURE Scales. Med Decis Making. 2019;39(6):673–80.
221. Forcino RC, Bustamante N, Thompson R, Percac-Lima S, Elwyn G, Pérez-Arechaederra D, et al. Developing and Pilot Testing a Spanish Translation of CollaboRATE for Use in the United States. PLoS ONE. 2016;11(12):e0168538.
222. De Las Cuevas C, Mundal I, Betancort M, Lara-Cabrera ML. Assessment of shared decision-making in community mental health

ARTICLE IN PRESS

care: Validation of the CollaboRATE. Int J Clin Health Psychol. 2020;20(3):262–70.
223. Forcino RC, Thygeson M, O'Malley AJ, Meinders MJ, Westert GP, Elwyn G. Measuring Patient-Reported Shared Decision-Making to Promote Performance Transparency and Value-Based Payment: Assessment of collaboRATE's Group-Level Reliability. J Patient Exp. 2020;7(5):742–8.
224. Barr PJ, Thompson R, Walsh T, Grande SW, Ozanne EM, Elwyn G. The psychometric properties of CollaboRATE: a fast and frugal patient-reported measure of the shared decision-making process. J Med Internet Res. 2014;16(1):e2.
225. Campbell J, Narayanan A, Burford B, Greco M. Validation of a multi-source feedback tool for use in general practice. Educ Prim Care. 2010;21(3):165–79.
226. Charalambous A. Validation and test-retest reliability of the Risser patient satisfaction scale in Cyprus. J Nurs Manag. 2010;18(1):61–9.
227. Charalambous A, Adamakidou T. Risser patient satisfaction scale: a validation study in Greek cancer patients. BMC Nurs. 2012;11:27.
228. Ventura MR, Fox RN, Corley MC, Mercurio SM. A patient satisfaction measure as a criterion to evaluate primary nursing. Nurs Res. 1982;31(4):226–30.
229. Biering P, Becker H, Calvin A, Grobe SJ. Casting light on the concept of patient satisfaction by studying the construct validity and the sensitivity of a questionnaire. Int J Health Care Qual Assur Inc Leadersh Health Serv. 2006;19(2-3):246–58.
230. Yellen E, Davis GC, Ricard R. The measurement of patient satisfaction. J Nurs Care Qual. 2002;16(4):23–9.
231. Kasper J, Hoffmann F, Heesen C, Köpke S, Geiger F. Completing the third person's perspective on patients' involvement in medical decision-making: approaching the full picture. Z Evid Fortbild Qual Gesundhwes. 2012;106(4):275–83.
232. Carryer J, Budge C, Hansen C, Gibbs K. Modifying the PACIC to assess provision of chronic illness care: an exploratory study with primary health care nurses. J Prim Health Care. 2010;2(2):118–23.

ARTICLE IN PRESS

233. Gugiu C, Coryn CL, Applegate B. Structure and measurement properties of the Patient Assessment of Chronic Illness Care instrument. J Eval Clin Pract. 2010;16(3):509–16.
234. Choi S-Y. The Effects of Leadership on Social Workers' Organizational Commitment: Focusing on Ethical Leadership, Organizational Trust, and Supervisor Trust. International Information Institute (Tokyo) Information. 2014;17(10 (A)):4771.
235. Asif M, Jameel A, Hussain A, Hwang J, Sahito N. Linking Transformational Leadership with Nurse-Assessed Adverse Patient Outcomes and the Quality of Care: Assessing the Role of Job Satisfaction and Structural Empowerment. Int J Environ Res Public Health. 2019;16(13).
236. Finn NK, Torres EM, Ehrhart MG, Roesch SC, Aarons GA. Cross-Validation of the Implementation Leadership Scale (ILS) in Child Welfare Service Organizations. Child Maltreat. 2016;21(3):250–5.
237. Lyon AR, Cook CR, Brown EC, Locke J, Davis C, Ehrhart M, et al. Assessing organizational implementation context in the education sector: confirmatory factor analysis of measures of implementation leadership, climate, and citizenship. Implement Sci. 2018;13(1):5.
238. Stogdill R. Manual for the leadership behaviour development questionnaire: form XII Columbus: Bureau of business research. Ohio: Ohio State University. 1963.
239. Carswell CM, Lio CH, Grant R, Klein MI, Clarke D, Seales WB, et al. Hands-free administration of subjective workload scales: acceptability in a surgical training environment. Appl Ergon. 2010;42(1):138–45.
240. Hoonakker P, Carayon P, Gurses A, Brown R, McGuire K, Khunlertkit A, et al. Measuring workload of icu nurses with a questionnaire survey: The nasa task load index (TLX). IIE Trans Healthc Syst Eng. 2011;1(2):131–43.
241. Mohamed R, Raman M, Anderson J, McLaughlin K, Rostom A, Coderre S. Validation of the National Aeronautics and Space Administration Task Load Index as a tool to evaluate the learning curve for endoscopy training. Can J Gastroenterol Hepatol. 2014;28(3):155–9.

ARTICLE IN PRESS

ARTICLE IN PRESS

242. Liljamo P, Kinnunen UM, Ohtonen P, Saranto K. Quality of nursing intensity data: inter-rater reliability of the patient classification after two decades in clinical use. J Adv Nurs. 2017;73(9):2248–59.
243. Liljamo P, Kinnunen UM, Saranto K. Healthcare professionals' views on the mutual consistency of the Finnish Classification of Nursing Interventions and the Oulu Patient Classification. Scand J Caring Sci. 2016;30(3):477–88.
244. van Oostveen CJ, Ubbink DT, Mens MA, Pompe EA, Vermeulen H. Pre-implementation studies of a workforce planning tool for nurse staffing and human resource management in university hospitals. J Nurs Manag. 2016;24(2):184–91.

ARTICLE IN PRESS