

Review Article

Instruments Measuring Integrated Care: A Systematic Review of Measurement Properties

MARY ANN C. BAUTISTA,*
MILAWATY NURJONO,* YEE WEI LIM,*
EZRA DESSERS,[†] and HUBERTUS JM VRIJHOEF*^{‡,§,||}

**Saw Swee Hock School of Public Health, National University of Singapore;*
[†]*Centre for Sociological Research, KU Leuven, Leuven, Belgium;* [‡]*National University Health System, Singapore;* [§]*Department of Patient and Care, University Hospital Maastricht, Maastricht, The Netherlands;* ^{||}*Vrije Universiteit Brussels, Brussels, Belgium*

Policy Points:

- Investigations on systematic methodologies for measuring integrated care should coincide with the growing interest in this field of research.
- A systematic review of instruments provides insights into integrated care measurement, including setting the research agenda for validating available instruments and informing the decision to develop new ones.
- This study is the first systematic review of instruments measuring integrated care with an evidence synthesis of the measurement properties.
- We found 209 index instruments measuring different constructs related to integrated care; the strength of evidence on the adequacy of the majority of their measurement properties remained largely unassessed.

Context: Integrated care is an important strategy for increasing health system performance. Despite its growing significance, detailed evidence on the measurement properties of integrated care instruments remains vague and limited. Our systematic review aims to provide evidence on the state of the art in measuring integrated care.

Methods: Our comprehensive systematic review framework builds on the Rainbow Model for Integrated Care (RMIC). We searched MEDLINE/PubMed for published articles on the measurement properties of instruments measuring integrated care and identified eligible articles using a standard set of selection

criteria. We assessed the methodological quality of every validation study reported using the COSMIN checklist and extracted data on study and instrument characteristics. We also evaluated the measurement properties of each examined instrument per validation study and provided a best evidence synthesis on the adequacy of measurement properties of the index instruments.

Findings: From the 300 eligible articles, we assessed the methodological quality of 379 validation studies from which we identified 209 index instruments measuring integrated care constructs. The majority of studies reported on instruments measuring constructs related to care integration (33%) and patient-centered care (49%); fewer studies measured care continuity/comprehensive care (15%) and care coordination/case management (3%). We mapped 84% of the measured constructs to the clinical integration domain of the RMIC, with fewer constructs related to the domains of professional (3.7%), organizational (3.4%), and functional (0.5%) integration. Only 8% of the instruments were mapped to a combination of domains; none were mapped exclusively to the system or normative integration domains. The majority of instruments were administered to either patients (60%) or health care providers (20%). Of the measurement properties, responsiveness (4%), measurement error (7%), and criterion (12%) and cross-cultural validity (14%) were less commonly reported. We found <50% of the validation studies to be of good or excellent quality for any of the measurement properties. Only a minority of index instruments showed strong evidence of positive findings for internal consistency (15%), content validity (19%), and structural validity (7%); with moderate evidence of positive findings for internal consistency (14%) and construct validity (14%).

Conclusions: Our results suggest that the quality of measurement properties of instruments measuring integrated care is in need of improvement with the less-studied constructs and domains to become part of newly developed instruments.

Keywords: integrated care, measurement instruments, quality of measurement properties, instrument validation, systematic review.

HEALTH SYSTEMS WORLDWIDE ARE IN THE MIDST OF A demographic and epidemiological transition marked by population aging¹⁻³ and an increasing burden of chronic diseases and disability.⁴ The synchronism of the change in demographics with the shift in disease patterns around the world inevitably impacts the health needs of populations. Health systems are compelled to make the necessary adjustments in order to continue supporting these needs. Along these lines, the World Health Organization (WHO) developed the

Innovative Care for Chronic Conditions (ICCC) Framework in 2002 to facilitate a new way of organizing health care systems. Expanding on the Model for Effective Chronic Illness Care,⁵ one of the guiding principles of the ICCC framework is building integrated health care to improve chronic disease management.⁶

This growing importance of integrated care draws attention to the need for systematically investigating how stakeholders interpret and measure integrated care. In this widely evolving field of research, pioneers have regarded integrated care as a principal strategy for improving patient care and increasing health system performance. These improvements are often measured in terms of enhanced quality of the patient-care experience, better health and well-being of communities, and reduced per capita health care costs.⁷⁻¹² Despite such significance, detailed evidence on the measurement properties of standardized and validated instruments that measure integrated care remains vague and limited. This limitation may be partly due to the similarly evolving methodologies for investigating instruments in this field of research.

Our systematic review aimed to provide evidence on the state of the art in measuring integrated care, including an assessment of the measurement properties of available instruments. Toward this end, we developed a comprehensive systematic review framework from the literature. We used this framework to identify the instruments that measure integrated care and the constructs used to describe the degree of integration. Finally, we synthesized the evidence on the quality of the studies included and the measurement properties of the integrated care instruments we identified. Exploring integrated care in this milieu potentially offers practical insights that could effectively advance the current understanding, design, implementation, and evaluation of integrated care.

We begin this paper with an overview of the concept of integrated care and its measurement. Key theoretical, conceptual, and measurement frameworks are discussed to lay the groundwork for presenting the results of the systematic review. We subsequently discuss the significance of gathering evidence on the measurement properties of instruments that measure integrated care. A background on the methodology of the systematic review and the psychometric evaluation of instruments is also presented. This will enable the understanding of the fundamentals in performing a systematic review of the measurement properties of instruments and highlight the unique contribution of this study.

Defining Integrated Care

The burgeoning of integrated care as a novel area of formal research and practice started around two decades ago.⁸ Over the years, relevant works have created awareness and stimulated greater interest in the meaning, rationale, implementation, and implications of integrated care.^{7-9,11,13-34} Although health policies dating back to the 1940s shaped more recent integrated care initiatives,³⁵ the imperative to remodel health systems may have also fueled the growing interest in integrated care. Health systems around the world have gradually evolved to better address the changing needs of the populations they serve.^{5,32} Recent health care reforms in the United States, for example, prompted the emergence of accountable care systems, one of the earlier conceptualizations of integrated service delivery.³⁶

An inherent challenge in relatively nascent fields such as integrated care is the surrounding conceptual ambiguity. Key stakeholders (ie, patients, health care providers, organization managers, and policymakers) have different perceptions of integrated care.^{8,11} In espousing a health system perspective, our systematic review adopted the following working definition of “integrated health care”: “the management and delivery of health services such that people receive a continuum of health promotion, health protection, and disease prevention services, as well as diagnosis, treatment, long-term care, rehabilitation, and palliative care services through the different levels and sites of care within the health system and according to their needs.”³⁷ Among a few alternative definitions, we adopted this definition from the Pan American Health Organization and the WHO European Office for Integrated Health Care Services because of its comprehensive scope. Further, as the leading agency for international health affairs, the WHO recommends strategies that are often modifiable within the specific context of its individual member states.³⁸ Hence, in using this definition, insights drawn from this paper can be potentially relevant to a wider international audience.

Key Theoretical and Conceptual Frameworks

The current literature contains a substantial amount of work synthesizing the key concepts of integrated care.^{8,11,16,24,29-31} Leutz previously

outlined the principles of integrating medical and social services to clarify arguments surrounding the early understanding of integration.^{7,9} However, the majority of these studies acknowledged that conceptual ambiguity remains a challenge. In fact, Minkman described integrated care as having various forms and lacking well-defined boundaries among its core concepts.³¹ Indeed, several terms have been used in different health systems across the world to denote integrated care. These include managed care (United States),³⁹ shared care (United Kingdom),¹⁶ transmural care (The Netherlands),⁴⁰ and other terms such as continuity of care, case management, patient-centered care, transitional care, and integrated delivery systems, to name a few.

There are various integrated care models across Western Europe (eg, United Kingdom,^{12-14,18,41} The Netherlands,^{13,32,42,43} Denmark and Finland,⁴⁴ Austria and Germany,³² and Sweden²⁷) and North America (eg, the United States^{7,9,39,45,46} and Canada²⁸). Examples include hospital-physician relationships, health maintenance organizations (HMOs),^{15,46} accountable care organizations (ACOs),^{39,47-49} the chronic care model,^{5,6} chains of care, and managed clinical networks.^{18,27,50}

Some of the widely held integrated care models and strategies emerged from the American health system. Kreindler and colleagues found that the unique social identities of individual organizations influence perceptions of implementers of integrated care (eg, ACOs).³⁹ The American model of HMOs is another often-cited example of integration.^{48,50} The HMO model, although based on a nonstandard definition, also contributed to generating some of the earliest concepts of integrated care.¹⁶ Lessons from successful HMOs (eg, Carle, Marshfield Clinic, Kaiser Permanente, Geisinger Health System)^{15,41,50} guided the design, implementation, and assessment of subsequent strategies for integrated care delivery.

Studies on integrated care models across European health systems also made substantial contributions in setting the foundations of integrated care. Disease management programs (The Netherlands and United Kingdom), chains of care (Sweden), and managed clinical networks (Scotland), for example, focus on meso-level integration that involves care provision to particular groups of patients and populations.⁴¹ Other known integrated care models across Europe are presented in greater detail in the literature.⁵¹

These above-mentioned models demonstrate how integrated care intrinsically relates to multidisciplinary components, objectives, and

varying perspectives. Other studies that attempted to surmount this conceptual challenge have recognized the dynamic and polymorphous nature of the integrated care concept.^{8,24} For example, Edgren viewed integrated health services as a complex adaptive system—an ever-changing organic entity involving various interrelated drivers.²² Similarly, Minkman introduced an empirically tested conceptual framework known as the Development Model for Integrated Care (DMIC).³¹ Although limited within the context of Dutch disease management programs, the DMIC also recognized integrated care as a complex, multilevel process having management and organizational components that are simultaneously at work.³³ Another study that used a structured framework to better understand integrated care, focused on the enabling mechanisms of integrated care for managing patients with chronic disease.⁵² Valentijn et al. also recently introduced the Rainbow Model of Integrated Care (RMIC).⁵³ As a comprehensive model, the RMIC supports a more coherent understanding and interpretation of the interactions among the complex dimensions of integrated care.³³

Measuring Integrated Care

The preceding discourse points to the considerable progress in constructing models and exploring strategies for developing and implementing integrated care. Thus, investigating its measurement can also benefit from adapting a similar structured approach especially because—akin to hitting a moving target—integrated care is a rapidly advancing field.

The often-discussed conceptual ambiguity of integrated care is a key deterrent in measuring the dimensions of integration. Moreover, instrument development itself is an iterative process that requires considerable effort and time to adequately capture the integrated care concept, which makes its measurement even more challenging. As a result, theoretical foundations and psychometric assessment of existing instruments have become limited.²⁰ Unlike more established research areas with clear-cut methods of investigation, measuring a dynamic and multifaceted system such as integrated care can be quite complex. Nevertheless, systematic measurement methods are vital in continuously developing the knowledge base of integrated care. Measurement is key in

identifying strategic areas for improving patient care and health system performance.

Existing reports on integrated care measurement range from systematic reviews to less-structured but comprehensive literature reviews of the constructs commonly measured; the measurement approaches and methodologies; and the actual instruments used in measuring constructs related to integrated care. Based on existing conceptual frameworks, previous studies measured various constructs that relate to several components of integrated care. The constructs commonly measured and described in scoping the literature include person/patient-centered care,^{54,55} continuity of care,^{56,57} care coordination,⁵⁸⁻⁶⁰ chronic disease management,⁶¹ quality measures in integrated care settings,⁶² integrated health care delivery,^{63,64} teamwork in health care settings,⁶⁵ and experiences of integrated care.^{66,67}

In terms of the measurement approaches, questionnaire surveys, registry data, and mixed data sources were identified as the most common methodologies in measuring integrated care.⁶⁴ Other models for integrated care measurement further characterized the measurement of integrated care according to the levels, degree, or continuum of integration. For example, Leutz described the 3 levels of integration in the context of integrated medical and social services (ie, linkage, coordination, and full integration).⁷ Based on this model, Minkman suggested that specific needs of service users determine the required level of integration. For instance, linkage of services would be sufficient for clients with less complex needs whereas those who have more complex needs would require full integration.^{68,69} In practice, however, integration involves a combination of these 3 levels of integration.⁹ Other research described a measurement model in terms of a continuum or cycle anchored by the 2 extremes of full segregation and full integration.^{17,30} Browne et al. also developed a comprehensive model for measuring multiple dimensions of integrated human service networks. These networks consist of multisector agencies working together to provide a continuum of services for individuals with complex needs.²⁰ Finally, Singer et al. proposed a framework using dimensions of coordination and patient-centeredness to measure the degree of patient care integration.²⁹

In terms of gathering evidence on the quality of measurement instruments, a recent systematic review focused on identifying available instruments for measuring care continuity.⁷⁰ Further, they assessed

the quality of the measurement properties of the available instruments. In so doing, the said review set a precedent in the systematic investigation of measurement properties of instruments in this field of research (vis-à-vis traditional instruments for health status measurement).

These studies on the relevant constructs and measurement models offer a glimpse of 2 main themes in integrated care measurement: (1) *what* is measured and (2) *how* it is measured. The current literature fairly indicates that integrated care measurement largely focuses on measuring individual aspects of integrated care. This suggests an atomistic rather than comprehensive understanding of integrated care measurement. Although the value of using a comprehensive tool for measuring integrated care remains to be justified, adopting a comprehensive framework for measurement provides the opportunity to identify other equally important aspects of integrated care that remain under-investigated. Furthermore, using a comprehensive instrument may also be relevant given the comprehensive scope of the definition of integrated care.

The current literature on integrated care measurement also points out the important distinction between the measurement of structure/process (ie, implementation and extent of achieving integration) and the measurement of outcomes (ie, evidence of effectiveness). Prior systematic reviews on health systems integration found that studies describing instruments that can measure both the processes and the outcomes of integrated care are limited,^{12,24,64} perhaps reasonably so, to avoid confounding by single-source bias. Further, a previous study maintained that capturing each dimension of an integrated system (ie, structural input, function, and output) would require a specific measurement instrument.²⁰ Although existing methods for measuring integrated care usually focus on the structures and processes (ie, at the organizational and administrative levels) rather than the outcomes (ie, impact),¹² triangulating these measures with clinical outcomes enhances the interpretability of their findings.⁷¹ In our systematic review, we deemed the measurement of structures and processes as more informative in describing the actual delivery (providers) and experience (users) of integrated care, compared with the measurement of outcomes. The literature on evaluating the quality of medical care⁷² supports this assumption, as we further explain in the discussion of the systematic review framework.

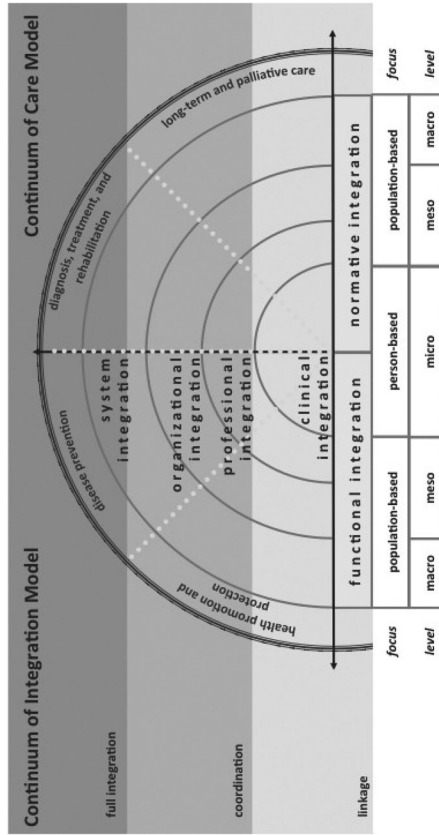
Systematic Review Framework

We developed a systematic review framework (Figure 1) to operationalize the concept and measurement of integrated care for our study and to enable the systematic evaluation of the instruments. This review framework was built on key studies on integrated care^{7,30,42,43,53} and the relevant publications from global health and research agencies.^{10,12,37,73,74} Based on our working definition, we put together key elements of the RMIC,^{42,43,53} the Institute of Medicine (IOM) continuum of care model,⁷⁵ and the continuum of integration model.^{17,30,53}

The RMIC sits at the core of our systematic review framework (see Figure 1). The RMIC is a comprehensive model that presents the 6 dimensions of integrated care: clinical, professional, organizational, system, functional, and normative integration. The first 4 dimensions operate at the micro, meso, and macro levels of the health system where integration presumably takes place. The last 2 dimensions (functional and normative integration) support the linkages across the health system levels and the other dimensions in the model. The consensus-based taxonomy of the key features of these dimensions is described elsewhere.^{42,43}

The rationale for applying the RMIC was twofold. First, the adoption of a comprehensive definition of integrated care warrants an equally comprehensive measurement framework. As such, the RMIC is one of the most comprehensive and tested models of integrated care. This unified conceptual framework can help address the ambiguity surrounding integrated care and guide subsequent investigations on the development of standards for its measurement. Second, the RMIC is based on the primary care tenets of person-focused and population-based care. Fundamentally, it uses primary care as a focal point to support integrated care. Integrated care offers to transfer the focus of care from high-cost hospitalizations to lower-cost ambulatory settings. In OECD countries, referral patterns appear to channel the entry of patients into the health care system via the primary care provider. Therefore, access to specialist care in these settings is often associated with a primary care referral. As gatekeepers, primary care providers are then increasingly important and potentially the best positioned to undertake care coordination, especially for patients with comorbidities.⁷⁶ In line with presenting the state of the art in measuring integrated care, it is only reasonable to adapt a framework that includes the key elements of primary care—a core component of integrated care.

Figure 1. Systematic Review Framework



Rainbow Model of Integrated Care

We also included the IOM continuum of care model in our framework to depict the service delivery component of integration as a continuum that covers health promotion, disease prevention, disease detection and management, and long-term care. Finally, to conceptualize the measurement of integrated care, our systematic review framework also adopted the continuum of integration model.^{17,30,53} This measurement model describes the degree of integration in terms of a continuum that spans the 2 extremes of full segregation and full integration; linkages, coordination, and cooperation lie between these 2 extremes.¹⁷

Finally, to reiterate an important assumption, our systematic review focused on the measurement of structures and processes that depict the experience of integrated care. This approach was based on examples from the literature^{20,30,58,63,77} that applied Donabedian's structure-process-outcome framework to evaluate the quality of medical care.⁷² The Donabedian model expresses important caveats in using outcomes in quality assessment. According to this model, outcome measures provide limited insights into the type and setting of the weaknesses or strengths they are supposed to indicate. While Donabedian recognized that outcomes remain as the ultimate indicators of the effectiveness and quality of care, examining the structure and process of care itself appears to better characterize the nature of the care provided and experienced. Moreover, enhancing the structures and processes is important in improving performance from a system perspective.¹⁴ This further underscores the value of measuring the structures and processes involved in integrated care—even though measuring the outcomes in terms of health system performance and patient health appears more relevant for other stakeholders.

Significance of the Systematic Review

When confronted with the task of selecting the appropriate instrument to measure integrated care, end users (ie, researchers, policymakers, and practitioners) could benefit from a comprehensive summary of instruments to guide their choice. The measurement properties of instruments are key considerations in making a well-informed decision. Measurement properties are specific attributes of an instrument that characterize the strengths and weaknesses of a test or measurement.⁷⁸ Instruments with good measurement properties are known to be more powerful

in obtaining strong conclusions and allow for better interpretation of findings.⁷⁹

In attempting to fill specific gaps in the literature, we conducted a systematic review of measurement properties of instruments measuring integrated care. To our knowledge, our study is the first of its kind to adopt a comprehensive systematic review framework that extensively considered construct and measurement models in the integrated care literature. Our systematic review aimed to provide new insights on integrated care measurement by answering the following research questions: (1) What instruments are currently available to measure integrated care? (2) What constructs were used to describe the degree of integration? and (3) What is the strength of evidence supporting the value of the measures reported (ie, the quality of the studies and the measurement properties of instruments)?

A systematic review of the measurement properties of instruments is designed to provide information that can support and justify the development of new instruments.⁸⁰ Based on our findings, the inclusion of the less-studied constructs and domains in new instruments may also be considered. More broadly, our review can contribute in setting the research agenda for further validating the available instruments and help introduce valuable insights in the field of integrated care and its measurement.

Methods

We developed and implemented a systematic review protocol (available upon request) based on published guidelines for best practice in systematic reviews of the measurement properties of instruments.⁸⁰⁻⁸⁴ We also adapted the recommended procedures from prior peer-reviewed studies on measurement properties and integrated care measurement.^{58,63,70}

Data Source and Search Strategy

We searched MEDLINE/PubMed in June 2014 and March 2015 (last updated search) and retrieved articles on the measurement properties of instruments measuring integrated care. One round of reference mining was performed on the eligible articles to capture additional relevant articles. Our systematic review framework, along with key papers on quality

| Component | Description | Example of Key Terms | Remarks |
|------------------------|---|---|--|
| Construct | Integrated care search terms | continuity of care, coordination of care, integrated care, patient-centered care, case management | Modified from Uijen et al. ⁷⁰ to fit the conceptual framework |
| Instrument | Search terms for the instrument of interest | questionnaires, instruments, measure, survey validation studies, | User-defined based on Terwee et al. ⁸¹ |
| Measurement properties | Measurement properties search filter | measurement properties, inter-rater | Directly applied from Terwee et al. ⁸¹ |
| Exclusion | Exclusion filter | editorial, interview, legal cases | Directly applied from Terwee et al. ⁸¹ |

assessment of the measurement properties of instruments, guided the search strategy development.^{70,81-84} We modified the search strategy from previous research⁷⁰ to correspond with our review framework before applying an additional search filter⁸³ for studies on the measurement properties of instruments. Key terms for the construct search included “continuity of care,” “coordination of care,” “integrated care,” “patient-centered care,” and “case management.” We summarized the key components of the final search strategy (online Appendix A) in Table 1.

Study Selection

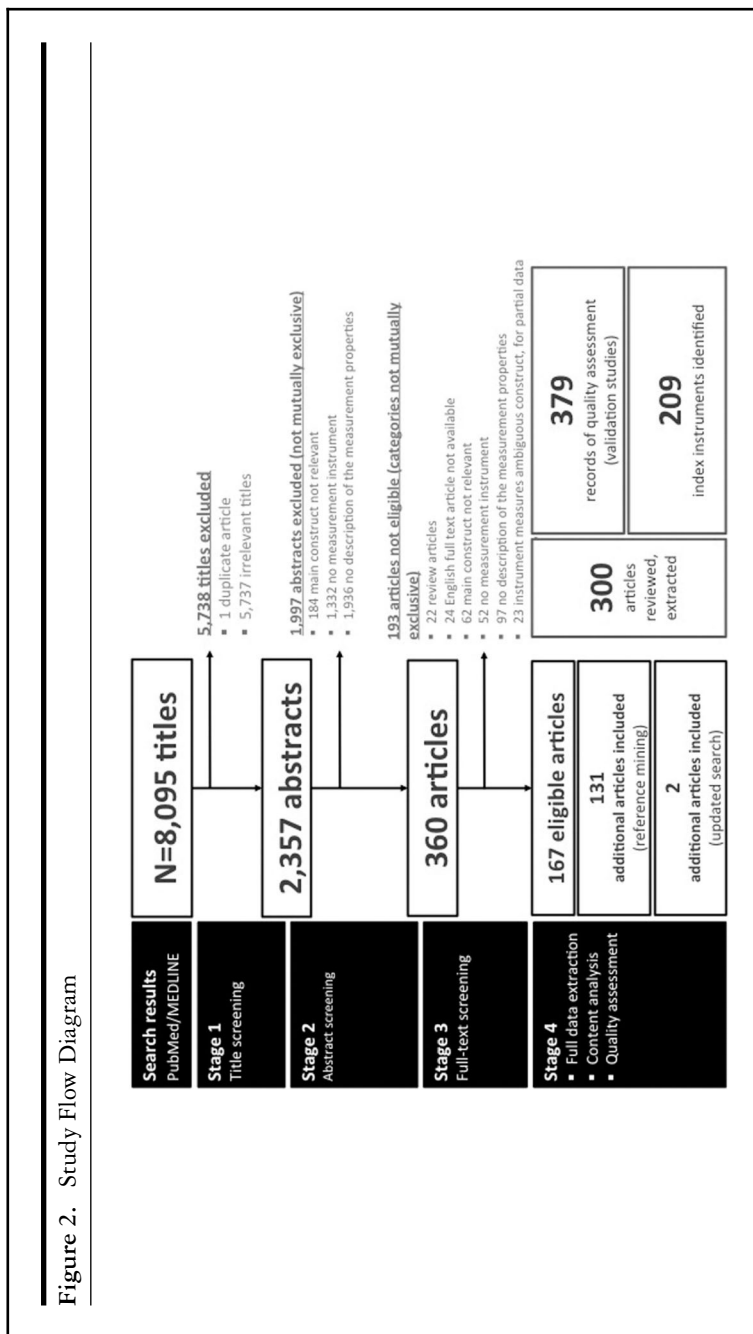
We selected eligible articles based on 4 main criteria: (1) availability of English full-text article, (2) measurement of constructs of integrated care as defined in the systematic review framework, (3) use of a relevant instrument or questionnaire, and (4) adequate description of the

development and/or validation of an instrument. Research metrics define a construct as “a characteristic or trait individuals possess to varying degrees that a test is designed to measure.”⁸⁵ Our review defined a relevant construct as a feature of the integrated care experience that an instrument intends to measure. Given our previous assumption that structures and processes better depict the care experience than outcomes do, our study focused on finding instruments measuring constructs related to the structure and/or process of integrated care delivery. Therefore, the second and third selection criteria effectively excluded instruments that measure the impact of integrated care rather than the actual care experience.

These 4 criteria were pilot tested on a sample of the search results prior to conducting the review. Thereafter, we independently screened the titles (MACB) (Stage 1) and abstracts (MACB and MN) (Stage 2) of articles from the main search results to identify records for full-text retrieval. MACB and MN discussed and resolved inconsistent screening results and consulted 2 more researchers (YWL and HJMV) for confirmation if necessary. We (MACB and MN) further screened the articles based on the full text (Stage 3) to select the articles for inclusion in the final review (Stage 4). With best effort, we applied a similar selection procedure in the reference mining. The study flow diagram that details the study selection process along with the final search results is presented in Figure 2.

Data Extraction

Using the standard data extraction form we constructed for our review, we abstracted the following data from the eligible articles: (1) basic article information (ie, author, title, journal name, year of publication, study eligibility); (2) validation study details (ie, design, objectives, setting, country); (3) description of respondents (ie, type, sample population, size, mean age, gender, disease status); (4) instrument details (ie, name, purpose, language, number of items, response scales, constructs purported to measure, constructs and domains of integrated care relevant to the conceptual framework); (5) details of instrument development (ie, item generation, refinement procedures, administration, scoring methods, theoretical basis); and (6) results of statistical analyses and measurement properties evaluated (ie, statistical methods, reported values for each measurement property). Index instruments that measure



integrated care constructs were also identified from the eligible articles. An index instrument is a principal instrument or questionnaire from which other versions have been developed for validation (eg, the Care Transition Measure⁸⁶ is an index instrument available in different versions and several languages).

Our review primarily aimed to identify instruments that measure integrated care and to describe the relevant constructs and domains the instruments used to measure integrated care. Hence, we later mapped the original constructs purportedly measured in the integrated care instruments to our systematic review framework. Based on our search strategy and systematic review framework, we recoded the original constructs into 4 categories of key constructs (Table 2). Care integration instruments, for example, included instruments that clearly specified measuring integrated care/care integration. Also classified under this key construct were instruments measuring combinations of multiple constructs that are relevant to our conceptual framework (ie, any combination of these constructs: continuity of care, coordination of care, integrated care, patient-centered care, case management, patient experience, chronic care, long-term care, and quality of care).

We also classified the instruments according to the relevant domains of integrated care using the 6 dimensions of the RMIC (ie, clinical, professional, organizational, system, normative, and functional integration). The instruments were grouped into the 6 dimensions based on the taxonomy of key features described by the original framework developers.⁴² These steps facilitated a thorough and structured examination of the instruments and the constructs they measured.

Quality Assessment of Studies and Instruments

The COnsensus-based Standards for the selection of health status Measurement INstruments (COSMIN) guided the quality assessment of the studies and the instruments we identified in this review.^{84,87-89} In quality assessment, there is an important distinction between the *quality of a study on the measurement properties* and the *quality of an instrument*.⁹⁰

Quality Assessment of Studies. The assessment of the methodological quality of the studies was performed using the 4-point scale COSMIN checklist.⁸⁴ This checklist is comparable to others that assess the quality of other types of studies included in a systematic review.⁸⁹ It includes an item list assessing whether a particular study satisfied the

Table 2. Instrument Categories According to Key Constructs of Integrated Care

| Key Construct | Description of Instruments |
|---|--|
| Care integration | <ul style="list-style-type: none"> • Clearly specified measuring integrated care/care integration • Measured any combination of multiple constructs relevant to integrated care as defined in our framework and original search strategy (ie, continuity of care, coordination of care, integrated care, patient-centered care, case management, patient experience within the integrated care framework, chronic care, long-term care, and quality of care) |
| Care continuity/ comprehensive care | <ul style="list-style-type: none"> • Clearly specified measuring care continuity/comprehensive care • Measured other components related to care continuity/comprehensive care (ie, care transition, handover, linkages) • Clearly specified measuring care coordination/coordination of care/case management |
| Care coordination/ case management | <ul style="list-style-type: none"> • Measured other components related to care coordination (ie, teamwork) • Clearly specified measuring patient-centered care • Measured other components related to patient-centered care (ie, patient–health care provider relationship, patient–doctor communication, shared decision making, trust, patient safety, empathy, unmet patient needs, information sharing, patient activation) |
| Patient-centered care | <ul style="list-style-type: none"> • Measured other components related to patient-centered care (ie, patient–health care provider relationship, patient–doctor communication, shared decision making, trust, patient safety, empathy, unmet patient needs, information sharing, patient activation) |

methodological standards for every measurement property it examined (Table 3). Hence, we rated a study either as poor, fair, good, or excellent for each item in the relevant measurement property. After rating each item, we applied a worst-score-counts algorithm to derive a COSMIN checklist quality score per measurement property examined. For

Table 3. Definition of Measurement Properties

| Measurement Property | Definition ^a |
|--|---|
| Reliability | |
| 1. Internal consistency | The degree/extent to which items in a (sub)scale are inter-correlated, thus measuring the same construct |
| 2. Reliability | The proportion of the total variance in the measurements due to true differences among patients |
| 3. Measurement error | The systematic and random error of a patient's score that is not attributed to true changes in the construct to be measured |
| Validity | |
| 4. Content validity | The degree to which the content of an instrument is an adequate reflection of the construct to be measured |
| 5. Structural validity | The degree to which the scores of an instrument are an adequate reflection of the dimensionality of the construct to be measured |
| 6. Construct validity (hypothesis testing) | The degree to which the scores of an instrument are consistent with hypotheses (eg, with regard to internal relationships, relationships to scores of other instruments, or differences between relevant groups) based on the assumption that the other instrument validly measures the construct |
| 7. Criterion validity ^b | The extent to which scores on a particular questionnaire relate to a gold standard |
| 8. Cross-cultural validity | The degree to which the performance of the items on a translated or culturally adapted health-related patient reported outcomes (HR-PRO) instrument are an adequate reflection of the performance of the items of the original version of the HR-PRO instrument |
| Responsiveness | |
| 9. Responsiveness | The ability of an instrument to detect change over time in the construct to be measured |

^aAdapted from Uijen et al.⁷⁰ and Mokkink et al.¹⁰⁷^bAdapted from Terwee et al.⁸¹

example, even if a study scored excellent in 5 of the 6 items for internal consistency, the 1 item rated poor would give the study a poor rating for this measurement property. Each validation study reported in every eligible article was counted as 1 study and assigned a unique identifier. The identifiers served to distinguish a validation study reported within a single eligible article. Each report of an instrument validation had a corresponding quality assessment of all the measurement properties considered (eg, a COSMIN quality score for internal consistency, or any other measurement property examined).

Quality Assessment of Instruments. The quality of an instrument is determined by its actual measurement properties (ie, statistical parameters reported for reliability, validity, etc). Guidelines on the quality criteria for good measurement properties are available from the literature.^{70,81} However, as of this writing, a well-defined benchmark for what qualifies as adequate measurement property remains to be established.⁸⁹ Nevertheless, it is important to assess the quality of a study (ie, COSMIN quality scores) before evaluating the evidence on the quality of any instrument it reports (ie, the actual results of the measurement properties reported).⁸⁹ We therefore assessed the quality of the studies and the quality of the measurement properties of instruments separately.

Data Analysis and Best Evidence Synthesis

We combined all data extraction records to facilitate iterative data verification steps throughout the subsequent data analysis and synthesis. We implemented a basic data management system using SPSS version 20 and Microsoft Office 2010 (eg, Word, Excel) for the descriptive statistical analyses and qualitative data synthesis, respectively. We then presented our main results with a descriptive summary of (1) the eligible articles (ie, article information, study details, and respondents), (2) the identified instruments and the constructs they measured, (3) the methodological quality assessment of each validation study per article (ie, COSMIN quality scores), and (4) the evaluation of the adequacy (ie, overall quality) of measurement properties per index instrument.

In evaluating the adequacy of measurement properties, we considered the evidence on the quality of studies and measurement properties per index instrument. It is possible to perform quantitative analysis by statistical pooling of the actual measurement properties, provided that the studies are of at least fair quality and homogeneous. Homogeneity

requires the studies to be adequately similar in terms of study characteristics and the results of the measurement properties.^{80,90} So far, however, there is limited research on established methods for quantitative analysis of data from studies on measurement properties.⁸⁰

Alternatively, qualitative analysis using best evidence synthesis allows drawing conclusions on the measurement properties when statistical pooling cannot be supported.⁸⁰ In this approach, 4 factors determine the adequacy of the measurement properties of each index instrument: (1) the number of validation studies, (2) the methodological quality of the studies, (3) the adequacy of the results of the measurement property (positive or negative), and (4) the consistency of results from several studies on the same instrument.^{70,80,81} The qualitative approach takes into account cases where different studies assessed the measurement properties of the same index instrument. In such cases, the adequacy of measurement properties of an index instrument was assessed from the results of different studies.

Of the two approaches, the qualitative approach was deemed more appropriate for our review. Thus, we summarized the evidence on the quality of measurement properties using the modified quality criteria for measurement properties (Table 4) and the criteria for the levels of evidence and overall assessment of measurement properties of instruments (Table 5). These criteria were adapted from relevant systematic reviews of measurement properties and those from the COSMIN developers.^{70,80,84} MACB applied these criteria to all index instruments ($n = 209$). To improve the validity of the overall assessment, 2 coauthors (YWL and HJMV) randomly checked a subsample of these assessments (60 instruments), while another (MN) independently applied the criteria to a random selection of instruments (45 instruments).

Results

Our main search strategy retrieved 8,095 articles in June 2014. After systematically screening titles, abstracts, and full-text papers, we included 167 of those articles in the full review. We included 131 additional articles in the full review after screening several potentially relevant articles that were identified in the reference mining. We expected to derive a number of additional articles from reference mining given the surrounding conceptual ambiguity of integrated care. Finally, our last updated

Table 4. Criteria for Rating the Adequacy of the Reported Measurement Properties

| Measurement Property | Reported Result | Quality Criteria ^a |
|-------------------------|-----------------|---|
| Reliability | | |
| 1. Internal consistency | + | (Sub)scale unidimensional OR Cronbach's alpha(s) ≥ 0.70 |
| | ? | Dimensionality not known OR Cronbach's alpha not determined |
| | - | (Sub)scale not unidimensional OR Cronbach's alpha(s) ≤ 0.70 |
| | 0 | Did not assess internal consistency |
| 2. Reliability | + | ICC/weighted Kappa ≥ 0.70 OR Pearson's r ≥ 0.80 |
| | ? | Neither ICC/weighted Kappa, nor Pearson's r determined |
| | - | ICC/weighted Kappa < 0.70 OR Pearson's r < 0.80 |
| | 0 | Did not assess reliability |
| 3. Measurement error | + | MIC $>$ SDC OR MIC outside the LOA |
| | ? | MIC not defined |
| | - | MIC \leq SDC OR MIC equals or inside LOA |
| | 0 | Did not assess measurement error |
| Validity | | |
| 4. Content validity | + | The target population considers all items in the questionnaire to be relevant AND considers the questionnaire to be complete |
| | ? | No target population involvement |
| | - | The target population considers items in the questionnaire to be irrelevant OR considers the questionnaire to be incomplete |
| | 0 | Did not assess content validity |

Continued

Table 4. *Continued*

| Measurement Property | Reported Result | Quality Criteria ^a |
|--|----------------------------|---|
| 5. Structural validity | + ? - 0 | Factors should explain at least 50% of the variance Explained variance not mentioned Factors explain < 50% of the variance Did not assess structural validity Correlation with an instrument measuring the same construct ≥ 0.50 OR at least 75% of the results are in accordance with the hypotheses AND correlation with related constructs is higher than with unrelated constructs |
| 6. Construct validity (hypothesis testing) | + ? | Solely correlations determined with unrelated constructs Correlation with an instrument measuring the same construct < 0.50 OR < 75% of the results are in accordance with the hypotheses OR correlation with related constructs is lower than with unrelated constructs Did not assess construct validity |
| 7. Criterion validity ^b | - 0 + ? - 0 | Convincing arguments that gold standard is "gold" AND correlation with gold standard > 0.70 No convincing arguments that gold standard is "gold" OR doubtful design or method Correlation with gold standard ≥ 0.70 , despite adequate design and method Did not assess criterion validity |

Continued

Table 4. *Continued*

| Measurement Property | Reported Result | Quality Criteria ^a |
|--------------------------|-----------------|---|
| Responsiveness | | |
| 8. Responsiveness | + | Correlation with an instrument measuring the same construct ≥ 0.50 OR at least 75% of the results are in accordance with the hypotheses OR AUC ≥ 0.70 AND correlation with related constructs is higher than with unrelated constructs |
| | ? | Solely correlations determined with unrelated constructs |
| | - | Correlation with an instrument measuring the same construct < 0.50 OR $< 75\%$ of the results are in accordance with the hypotheses OR AUC < 0.70 OR correlation with related constructs is lower than with unrelated constructs |
| | 0 | Did not assess responsiveness |

Abbreviations: AUC, area under the curve; ICC, intra-class correlation coefficient; LOA, limits of agreement; MIC, minimal important change; SD, standard deviation; SDC, smallest detectable change.

^aThese criteria were adapted from Uijen et al.⁷⁰ unless stated otherwise. The relevant criterion for cross-cultural validity is not available from the current literature; this precluded rating the quality/adequacy of cross-cultural validity property in the review.

^bAdapted from Terwee et al.⁸¹

Table 5. Criteria for the Level of Evidence and Overall Assessment of Measurement Properties

| Criteria ^a | Overall Assessment | Level of Evidence |
|---|--------------------|-------------------|
| Consistent findings in multiple studies of good methodological quality OR in 1 study of excellent methodological quality | + + + or - - - | Strong |
| Consistent findings in multiple studies of fair methodological quality OR in 1 study of good methodological quality | + + or - - | Moderate |
| 1 study of fair methodological quality | + or - | Limited |
| Conflicting findings from multiple studies | +/- | Conflicting |
| Only studies of poor methodological quality OR only indeterminate results from multiple studies regardless of methodological quality | ? | Unknown |
| Measurement property not assessed | 0 | Not assessed |

^aAdapted from Uijen et al.⁷⁰

search in March 2015 yielded 2 more articles. Our study selection process amounted to 300 unique articles for inclusion in the full review.

We identified 209 index instruments measuring a relevant integrated care construct. We assigned a unique identifier to 379 instrument validation studies reported in the eligible articles, each with a corresponding quality assessment of measurement properties. We presented consumer tables in supplementary files (available online) to provide an overview of the key characteristics of the eligible articles (Online Appendix B); the identified instruments measuring integrated care in these articles (Online Appendix C); the methodological quality assessment of each

validation study (Online Appendix D); and the adequacy of measurement properties per index instrument (Online Appendix E). We structured the following summary of results according to these major themes. In addition, we provided a list of instruments that have evidence of adequate measurement properties (Online Appendix F) and another list of instruments grouped by domain (Online Appendix G) as a quick reference catalogue of instruments (see online notes on appendices).

Eligible Articles Included in the Full Review

The number of relevant publications on instruments measuring integrated care-related constructs showed an increasing trend over the years (Figure 3). The gradual annual increase in published articles started in 1998, peaking to a high of approximately 30 publications in 2013. The United States (38%), Canada (13%), the Netherlands (10%), the United Kingdom (10%), and Australia (6%) were the top 5 countries that contributed to the literature on the measurement properties of integrated care instruments since 1988, with a few contributions from Asia and Africa (Figure 4). Of the 300 articles included, nearly 60% were published in either a medical specialty (33%) or health services research journal (23%). Other common journal subject areas included nursing (12%), mental health and substance abuse (11%), and cancer and patient education (4% each), with a few other articles (<4%) from journals that focused on a range of areas, including clinical epidemiology, dental health, and service industry management.

Instruments Measuring Integrated Care

Constructs and Domains. Table 6 presents a descriptive summary of the general and construct-specific characteristics of instruments identified per validation study ($n = 379$). We initially determined the main constructs the instruments originally intended to measure. More than 50% of the instruments purported to measure patient-centered care. After recoding these constructs to match the key constructs in our framework, the majority of the instruments measured care integration (33%) and patient-centered care (49%). Fewer instruments measured constructs of care continuity/comprehensive care (15%) and care coordination/case management (3%). Furthermore, we mapped the majority of constructs

Figure 3. Number of Articles Published per Year from 1988 to 2015

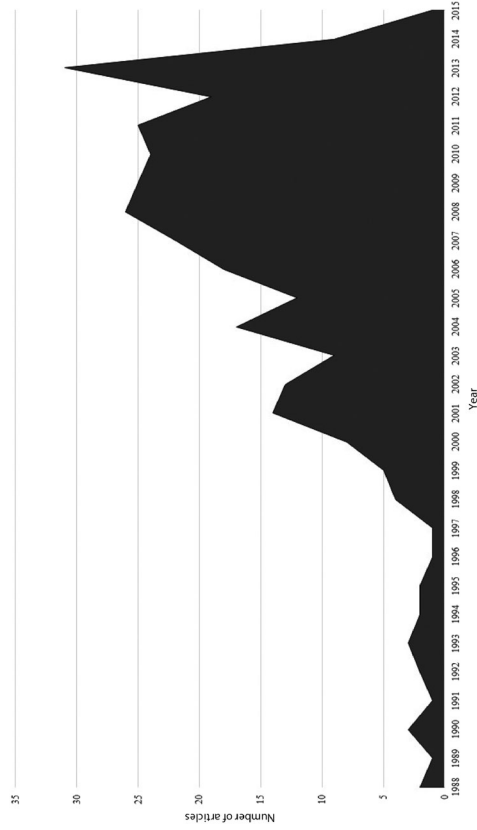


Figure 4. Distribution of Articles by Country of Publication

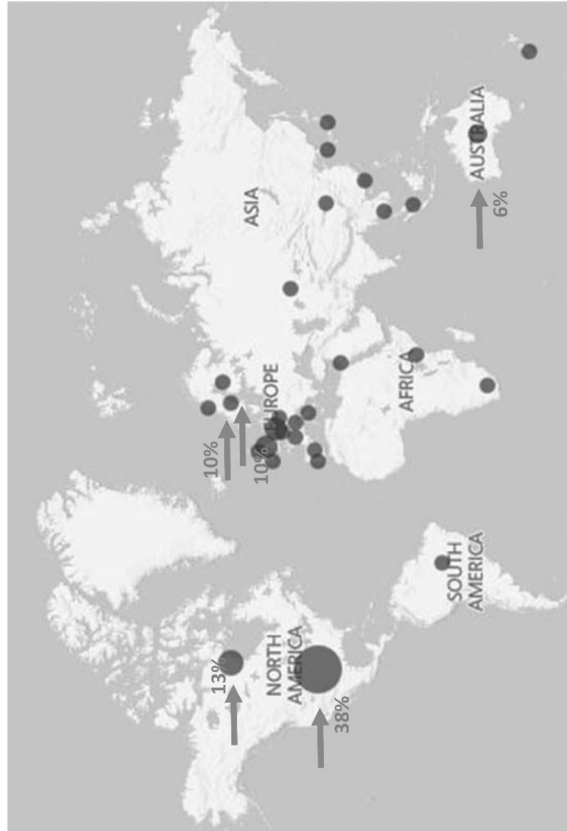


Table 6. Characteristics of Instruments Measuring Integrated Care Constructs Identified in the Review (379 validation studies of 209 index instruments)

| General Characteristics | |
|---|--------------|
| Type of respondent | n (%) |
| Patient themselves | 228 (60) |
| Informal caregiver themselves | 26 (7) |
| Patients (proxy) | 1 (<1) |
| Health care providers | 77 (20) |
| Patients and informal caregivers | 9 (2) |
| Patients and health care providers | 10 (3) |
| Other ^a | 28 (7) |
| Health status of respondent | |
| Healthy | 40 (11) |
| With disease | 219 (58) |
| Mixed | 23 (6) |
| Not applicable ^c | 66 (17) |
| Not reported | 31 (8) |
| Disease category (n = 219) | |
| Non-chronic condition | 3 (1) |
| Single chronic condition | 36 (16) |
| Multiple chronic conditions | 46 (21) |
| Cancer | 30 (14) |
| Mental and behavioral disorders | 35 (16) |
| Other nonspecific conditions ^d | 36 (16) |
| Not reported | 33 (15) |
| Setting (not mutually exclusive) | |
| General population | 39 (10) |
| Primary care | 148 (39) |
| Secondary care | 121 (32) |
| Specialist care | 127 (34) |
| Community services | 47 (12) |
| Home-based care | 20 (5) |
| Social services | 1 (<1) |
| Nursing homes | 16 (4) |
| Pharmacies | 2 (<1) |
| Mode of administration | |
| Face-to-face interview | 40 (11) |
| Tele-interview | 20 (5) |
| Self-administered (PAPI, postal) | 235 (62) |
| Self-administered (CAPI, electronic) | 13 (3) |

Continued

Table 6. *Continued*

| | |
|--|--------------|
| Combination | 26 (7) |
| Other ^b | 30 (8) |
| Not reported | 15 (4) |
| Language administered in | |
| English | 249 (66) |
| Dutch | 32 (8) |
| Swedish | 13 (3) |
| Spanish | 6 (2) |
| Chinese | 5 (1) |
| Finnish | 5 (1) |
| German | 4 (1) |
| Korean | 4 (1) |
| Norwegian | 4 (1) |
| Thai | 3 (<1) |
| Portuguese | 3 (<1) |
| French | 2 (<1) |
| Japanese | 2 (<1) |
| Other ^c | 5 (1) |
| Multiple languages (at least 2) | 33 (9) |
| Not reported | 9 (2) |
| Construct-specific Characteristics | |
| Constructs purported to measure | n (%) |
| Care integration | 25 (7) |
| Care continuity | 51 (14) |
| Care coordination | 26 (7) |
| Case management | 3 (<1) |
| Patient-centered care | 216 (57) |
| Comprehensive care | 5 (1.3) |
| Patient satisfaction | 5 (1.3) |
| Chronic care | 10 (3) |
| Multiple constructs | 28 (7) |
| Other ^g | 10 (3) |
| Constructs measured (mapped to the framework) | |
| Care integration | 124 (33) |
| Care continuity/comprehensive care | 58 (15) |
| Care coordination/case management | 12 (3) |
| Patient-centered care | 185 (49) |
| Domain classification (mapped to the framework)^h | |
| Clinical integration | 318 (84) |
| Professional integration | 14 (4) |

Continued

Table 6. *Continued*

| | |
|----------------------------|--------|
| Organizational integration | 13 (3) |
| System integration | 0 (0) |
| Functional integration | 2 (<1) |
| Normative integration | 0 (0) |
| Combination | 32 (8) |

Abbreviations: CAPI, computer-assisted personal interviewing; PAPI, paper and pencil interviewing; proxy, an informal caregiver or another representative responds on behalf of the patient.

^aDirect observer or rater, verbal coding, organizations, researchers, students.

^bDirect observer or rater, checklist.

^cRespondents were not patients.

^dOnly described as patients.

^eItalian, South African, Hebrew, Arabic, and Greek.

^fRehabilitation center, university/teaching institution/medical schools, intermediate care facilities, integrated care organizations, health education facilities, long-term care facilities, and dental facilities.

^gClinical risk management, cultural competency, geriatric care environment, quality of care, care processes, organizational access, etc.

^hNormative and system integration domains were always assessed in combination with the other domains.

to the clinical integration domain (84%), with fewer constructs related to the professional (3.7%), organizational (3.4%), and functional (0.5%) integration domains. None of the instruments exclusively assessed system or normative integration, and only 8% of the instruments measured a combination of any of the domains.

We also identified a few ambiguous instruments ($n = 47$). An ambiguous instrument purported to measure a construct that is not relevant to integrated care (as defined in our framework and original search strategy), but used key constructs of integrated care as indicators for the main construct it intended to measure. For instance, some patient satisfaction instruments measured patient satisfaction in terms of integrated care constructs (eg, satisfaction with continuity of care, process of care, etc). Hence, these instruments had components, dimensions, or subscales related to our definition of integrated care, but did not aim to directly measure integrated care. We excluded these instruments in the quality assessment to consistently adhere to our conceptual framework.

In 5 index instruments, we found different studies⁹¹⁻¹⁰² using similar instruments to measure 2 different constructs (details in

Appendix C). For example, different studies used the Primary Care Assessment Survey to measure different constructs such as interpersonal communication¹⁰²; accessibility, relational continuity, interpersonal processes of care, respectfulness, and management continuity¹⁰¹; and organizational access.¹⁰⁰ Likewise, a study used the Components of Primary Care Index to measure key aspects of primary care⁹⁹ and another for measuring comprehensive care.⁹⁸ The Consumer Quality Index was also used to measure quality of care in general and in disease-specific settings,⁹²⁻⁹⁷ while another study used it to measure patient experience in long-term care.⁹¹

Other Instrument Characteristics. The majority of instruments were administered to either patients (60%) or health care providers (20%). Other types of respondents included informal caregivers, proxies to patients, direct observers, or a combination of these types. More than half of the instruments involved patients with disease (58%), which ranged from non-chronic conditions, multiple chronic conditions, and cancers, to mental and behavioral disorders. Chronic conditions included, among others, injuries and conditions requiring rehabilitation, chronic infectious and noninfectious diseases, diabetes, hypertension, arthritis, heart conditions, and asthma. Other conditions designated as “nonspecific conditions” included frailty, poor health, hospitalized patients, and other common diseases. Cancer was included in a separate category to distinguish it from other chronic diseases. Historically, cancer is associated with an acute fatal illness, but recent developments in cancer therapy have changed the course and management of the disease.¹⁰³ Some cancer types are now being managed as a chronic disease. Because our systematic review did not distinguish between the specific cancer types, we did not elect to simply combine cancer with other chronic diseases.

Traditional self-administration using paper and pencil was the most common mode of administration (62%) followed by face-to-face interviews (11%), direct observation checklists (8%), and tele-interviews (5%). A few instruments used a combination of traditional and computer-assisted self-administered instruments as well as face-to-face and tele-interviews (8%). The instruments measured integrated care constructs in various settings, but mainly in primary (39%), secondary (32%), and specialist care (34%). Although the instruments were developed in various languages, the majority were in English (66%), with relatively fewer instruments in Dutch (8%) or Swedish (3%).

Type of Respondents. Finally, we cross-tabulated the constructs measured with the type of respondents to further characterize the instruments we identified. In general, instruments measuring integrated care constructs were largely administered to patients, except for instruments that measure care coordination and case management, which largely targeted health care providers (Figure 5). All key constructs were consistently measured in either patients or health care providers, but very few in both (eg, 8% of instruments measured care coordination/case management). While the proportions of instruments administered to informal caregivers were noteworthy in those that measure care integration (64%) and patient-centered care (58%), the use of patient proxies in instrument validation was largely unremarkable (only 0.8% of the care integration instruments). Examining the instruments in terms of the domains and type of respondents revealed a similar pattern (Figure 6). Most instruments classified under each domain were administered to either patients or health care providers. For example, instruments measuring clinical integration constructs were administered to patients (60%) or health care providers (15%). Moreover, the instruments generally targeted patient respondents more than the health care providers, except in the professional integration domain in which 57% of instruments were administered to health care providers.

Quality of Studies: The COSMIN Quality Scores

Table 7 presents a summary of the measurement properties of instruments evaluated in the validation studies ($n = 379$). At least half of the studies independently assessed the following measurement properties: internal consistency (91%), content validity (63%), structural validity (82%), and construct validity (55%). Cross-cultural (14%) and criterion (12%) validity, measurement error (7%), and responsiveness (4%) were less commonly evaluated.

For each validation study, the COSMIN quality scores obtained per measurement property indicate the quality of the study for the measurement property reported (Figure 7). We found that <50% of the validation studies were of good or excellent quality for any of the measurement properties. Studies that assessed internal consistency were mostly of fair (32%) or excellent (22%) quality. Structural validity showed a similar

Figure 5. Instrument Validation According to the Constructs and Type of Respondents

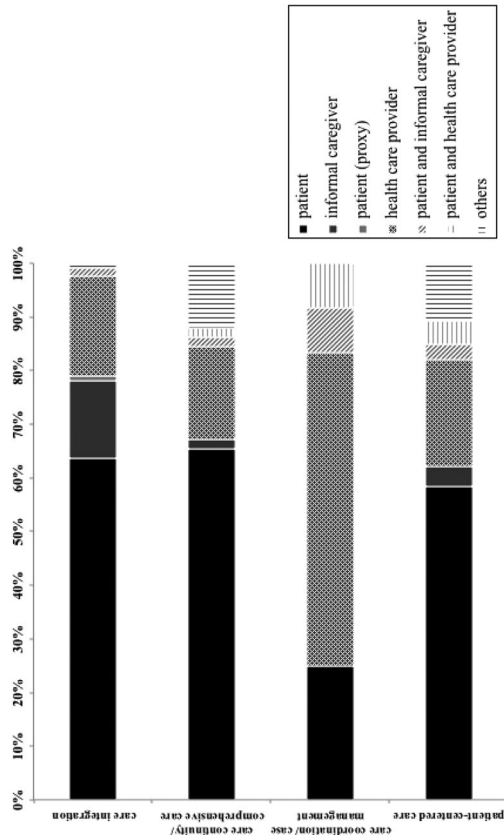


Figure 6. Instrument Validation According to the Domains and Type of Respondents

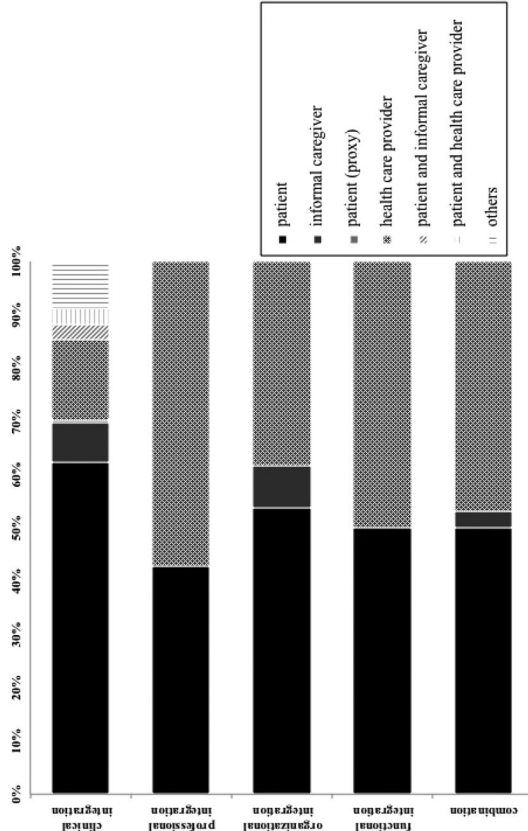


Table 7. Summary of Measurement Properties Assessed in Records of Instrument Validation (n = 379)

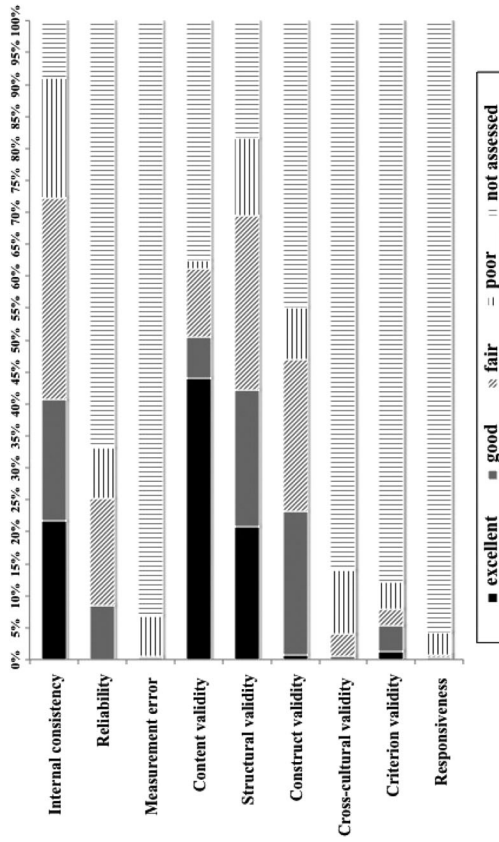
| Measurement Property | n (%) |
|---|----------|
| Internal consistency | 345 (91) |
| Reliability | 126 (33) |
| Measurement error | 26 (7) |
| Content validity | 237 (63) |
| Structural validity | 309 (82) |
| Construct validity (hypothesis testing) | 209 (55) |
| Cross-cultural validity | 53 (14) |
| Criterion validity | 46 (12) |
| Responsiveness | 16 (4) |

distribution where a larger proportion of studies were of fair (27%) or excellent (21%) quality. Content validity showed the highest percentage of excellent quality studies (44%), while internal consistency had the greatest percentage of poor quality studies (19%).

Quality of Instruments: The Best Evidence Synthesis

After separate assessments of the quality of studies (ie, COSMIN checklist) and the quality of instruments (ie, results of measurement properties reported), we combined the results from the 2 components to obtain the level of evidence and adequacy rating for each index instrument. Figure 8 shows the percentage of index instruments (n = 209) according to the level of evidence and adequacy of measurement properties. Consistent with the observed trends in each component assessed, at least 60% of the index instruments (n = 209) evaluated the following measurement properties: internal consistency (90%), content validity (78%), structural validity (74%), and construct validity (65%). The level of evidence for measurement error and responsiveness was largely unassessed, if not of unknown level. Correspondingly, the level of evidence for a considerable percentage of index instruments was unknown for internal consistency (24%), content validity (43%), structural validity (32%), and construct validity (17%). Nevertheless, a number of

Figure 7. COSMIN Scores: Methodological Quality of Studies per Measurement Property



index instruments showed strong evidence of positive findings for internal consistency (15%), content validity (19%), and structural validity (7%); with a few others showing moderate evidence of positive findings for internal consistency (14%) and construct validity (16%). A few index instruments showed strong evidence of negative findings for internal consistency (1%), reliability (0.5%), and structural validity (3%).

Discussion

Our systematic review examined studies on the measurement properties of instruments measuring integrated care. The study findings provide a comprehensive picture of important developments in integrated care measurement by presenting the currently available instruments and the relevant constructs and domains measured. We also summarized the strength of evidence for the adequacy of the measurement properties of these instruments. In the discussion, we further examine the implications of these findings within the context of the international literature and highlight the important research gaps addressed by the study.

Global Integrated Care Initiatives

The increasing number of relevant studies in our review underscores the growing interest in integrated care and its measurement. Although the majority of these studies involved work from Western Europe and North America, finding a few studies from Africa and Asia-Pacific indicates the potential for expanding integrated care initiatives across the world. While adopting the design, implementation, and evaluation of integrated care is by no means straightforward, it is certainly worth exploring how other regions can learn from the early models of integrated care delivery and its measurement. After all, the use of integrated care as an important strategy for increasing health system performance should benefit as many health systems as possible. However, we must also recognize that the timing and relevance of implementing integrated care may differ for each health system.

Available Instruments

Our search strategy identified 209 index instruments that used different constructs to measure integrated care. Compared with prior

systematic reviews on instruments,^{58,63,64,70} the large number of instruments we identified in our review is consistent with the growing interest in integrated care described in the literature. Conversely, the paradigm shift toward chronic disease management and the resulting health care reforms¹⁻⁴ also may have driven the increasing relevance of integrated care. The large number of instruments also underscores the need to explore and develop more structured methodologies in the metrics of integrated care. Observing a considerable number of validated instruments (39%) tested in the primary care setting is consistent with our use of a framework that is based on primary care (eg, RMIC).⁵³ In our review, we also found a few studies⁹¹⁻¹⁰² where similar instruments were used to measure different constructs of integrated care.⁹⁸⁻¹⁰² This further demonstrates the need for a more standard process of selecting appropriate instruments to measure a particular construct.

Clearly, the integrated care instruments we identified measured varied constructs. This finding can be a direct consequence of our comprehensive framework designed to capture any construct relevant to the integrated care experience. In addition, our review included studies from as early as the 1980s. Such studies predate the recently developed taxonomy and conceptual framework of integrated care we used, which may have contributed to the nonspecific conceptualization of integrated care. The variation in the constructs measured remains consistent with previous accounts that described the multidimensionality of integrated care.^{8,11} Moreover, the different constructs and dimensions measured in the instruments we identified can offer insights into the state of integrated care initiatives in the health systems where these instruments were developed. We venture this view given the importance of the “context of use” in instrument development and validation. Hence, the measurement model that is relevant to one health system may not necessarily work for another. For example, providers that focus on acute care services would be interested in evaluating structures and processes that are fundamentally different from those that have adopted a chronic care model (eg, integrating chronic disease management in primary care services). In this regard, the exclusion of the ambiguous instruments in our study emphasizes our focus on identifying instruments that measure the structure and process—rather than the outcomes—of integrated care. This important distinction underscores the importance of focusing on organizational structures and processes in improving overall performance of the system.¹⁴

Mapping the constructs purportedly measured by the instruments to our systematic review framework provided additional practical information. We categorized the majority of the instruments under patient-centered care and care integration, with fewer instruments measuring care continuity/comprehensive care and care coordination/case management. The relative abundance of instruments measuring patient-centered care is consistent with the majority of these instruments (60%) targeting patient respondents. These findings suggest the growing relevance of patients in assuming a more active role in the health system. Conversely, the findings may also indicate that health systems have started to recognize the importance of engaging patients in improving the health system.

In terms of the domains, we mapped only a few of the constructs to the professional, organizational, and functional integration domains, and we found that none of the instruments measured constructs specifically related to the normative or system integration domains. It is likely that the relevant constructs for these domains may be more deeply underlying (eg, organizational policies, value system),⁴² which could limit efforts to measure these areas. Although we mapped the constructs to combinations of these domains in some instruments, none of the instruments measured all the relevant constructs and domains of integrated care as defined in our review framework.

These results draw attention to the specific dimensions of integrated care that remain under-investigated. Other equally important constructs (ie, care continuity/comprehensive care and care coordination/case management) and domains (ie, system and normative integration) were not extensively examined. Further, we did not find one comprehensive instrument that measured all the relevant dimensions of integrated care. The fundamental structure of a health system may influence the decision to measure parts rather than the whole of integrated care. It is also likely that some health systems choose to focus on specific aspects of integrated care—while dispensing with other aspects—given their respective priorities and prevailing needs, goals, and standards in health care delivery. For example, a study on instrument development involving a health system that aims to strengthen both person-focused and population-based care may elect to develop and validate an instrument that captures all the dimensions of integrated care as depicted in our framework. Fundamentally, a partial understanding of integrated care may be suitable given the proper context and objective of measurement.

However, focusing on specific aspects of integrated care without the unifying elements of a comprehensive framework could further contribute to the conceptual ambiguity that has persistently plagued investigations in integrated care. It is also important to note the possibility that the capacity of standard measurement instruments may be limited and insufficient to adequately capture some dimensions of integrated care. Other data sources and measurement approaches (eg, registry data, mixed data sources, policy documents) may be needed to measure some dimensions.

While these findings clearly point to a measurement gap, they still imply that partial integrated care efforts are making headway. Finding studies that only measured certain dimensions of integrated care is still a good indication of the measurement of ongoing integrated care initiatives. Whether a comprehensive instrument is needed to provide a better measure of integrated care remains an outstanding question⁶³ and one that is beyond the scope of our study. Nevertheless, our systematic summary of the available instruments and their measurement properties provides a good starting point for subsequent investigations. Future studies should consider identifying and synthesizing all the necessary information that can provide sufficient evidence of the need for a comprehensive instrument. Doing so would entail a series of validity testing of the available comprehensive instruments.

Altogether, these findings on integrated care instruments reinforce assertions from previous studies that point to the lack of a standard definition and a unified framework for measuring integrated care. The complexity in operationalizing the integrated care concept has ever since rendered the task of measuring it even more challenging. Nevertheless, it is imperative to develop valid measures of integration across health care systems.¹⁶

Quality of Instruments

Besides identifying integrated care instruments and the constructs they measured, our systematic review also attempted to assess the quality of studies and the quality of instruments measuring integrated care. Applying the COSMIN checklist to these instruments showed that internal consistency, structural validity, content validity, and construct validity were the most commonly assessed measurement properties; whereas responsiveness, measurement error, and criterion and cross-cultural validity were less commonly evaluated. Less than 50% of the validation

studies were rated to be of good or excellent quality for any of the measurement properties. Only 19% of index instruments showed strong to moderate evidence of positive findings for some measurement properties (ie, internal consistency, content validity, structural validity, and construct validity). However, the strength of evidence on the adequacy of most of the measurement properties remained largely unassessed (see Figure 8). This is mainly due to having only poor quality studies or only indeterminate findings on the measurement properties of the instruments. As proffered in a previous systematic review of measurement properties of care continuity instruments,⁷⁰ these findings on the levels of evidence do not suggest that the available instruments are of poor quality. The results only highlight the need for more high-quality studies that can adequately assess the measurement properties and, ultimately, the instrument quality.

The latter findings further position our review in the context of the broad integrated care literature. Assessing the quality of studies and measurement properties of integrated care instruments is important in refining the knowledge base in this field. Other systematic reviews on the instruments measuring integrated care concepts^{58,63} did not include the quality assessment of the measurement properties of the instruments. Uijen and colleagues, however, provided a useful guide for our own work, as they closely examined the measurement properties of questionnaires measuring continuity of care.⁷⁰ Moreover, the pervasiveness of using questionnaires in measuring integrated care⁶⁴ suggests the need for a thorough validation of already existing measurement instruments across different settings.^{63,77} The results of our systematic review can facilitate setting the research agenda for validating the currently available instruments and informing the decision to develop new ones.⁸⁰ In doing this study, other research areas may take an interest in exploring the value of a systematic review of the measurement properties of instruments. And as more reviews of this kind are conducted, the methodologies for systematic reviews of measurement properties can be improved.

Limitations and Strengths of the Systematic Review

It is important to consider these findings in light of the limitations of the review. First, our comprehensive systematic review framework resulted in a heterogeneous pool of studies and instruments. This

heterogeneity in the study design characteristics and the statistical results of the measurement properties precluded more advanced quantitative analysis via statistical pooling. A quantitative approach could have improved the interpretability of the findings at a level sufficient to support recommendations for selecting the best instrument. Moreover, given the number of studies included in our review, we resolved to take the data from the studies at face value without verifying the data with the primary study investigators. Although we devised a basic data management system to maintain the validity of data extraction, the reporting of results in the primary studies also limited the quality of data extracted.

Second, we did not search the gray literature, which may contain relevant instruments. Considering the scope of our review, we anticipated a substantial number of relevant articles, which inevitably had an impact on the feasibility of searching multiple databases. Nevertheless, limiting the search to MEDLINE/PubMed is based on a reasonable assumption that the majority of relevant publications would have been indexed in PubMed, rather than in other databases that are recommended for clinical research and other sub-branches of medicine. Experts in this field recommend using at least MEDLINE and Embase.⁸⁰ Including Embase in our search, however, may not have provided additional relevant articles given that the database focuses more on drugs and pharmacology.¹⁰⁴ Similarly, searching Embase would yield a greater number of case reports,¹⁰⁵ which is an article type that is not relevant to our selection criteria.

Third, excluding non-English articles may have introduced language bias. We expect to minimize the potential influence of such a bias in maintaining that English is the standard language in journals that publish works on integrated care. Furthermore, there is no evidence that the exclusion of languages other than English introduces any systematic bias in reviews.¹⁰⁶ Finally, unlike conventional systematic reviews, we could not assess the impact of publication bias in the absence of standard methods of bias assessment for systematic reviews of measurement properties.⁸⁰

Aside from these limitations, a systematic review of the measurement properties of instruments also presents general methodological challenges and learning opportunities. This type of systematic review has multiple outcome measures (ie, constructs and measurement properties) that may not be consistently evaluated in all the studies. Conversely, the evidence for one measurement property may come from multiple

studies; each measurement property would then require its own set of data synthesis. Nevertheless, we put forth our best effort to present the quality of studies and measurement properties of instruments that would allow for a comprehensive overview of how integrated care is measured. Further, our data synthesis focused on the qualitative approach used in previous systematic reviews.^{58,63,70,81} This approach permitted the best possible synthesis of evidence that accounted for the between-study similarities, the quality of each study, and the consistency of measurement properties reported across different studies.⁸¹ Although the COSMIN checklist was originally developed for the assessment of instruments that intend to measure patient-reported health outcomes,¹⁰⁷ it was deemed an appropriate tool for assessing the quality of studies on other types of instruments.⁸⁰

Despite these limitations, a key strength of our review is the completion of the formidable task of presenting a comprehensive picture of the integrated care measurement landscape. Pursuing maximum effort in capturing and synthesizing data on an evolving research area is challenging. Still, the insights gained from our review findings outweigh the challenges that had to be overcome. These insights, however, are only as sound as the supporting framework for understanding and measuring integrated care. Hence, we presented our findings within a framework that combined core elements of the IOM continuum of care model, the continuum of integration model, and the RMIC, which is central to the current understanding of integrated care.⁵³ We identified a large number of studies on instruments measuring a range of different integrated care concepts. Moreover, our review included studies involving patients, informal caregivers, and health care providers that represent important groups of stakeholders in integrated care. We also exploited a rich dataset of instruments to provide potential end users with organized and structured information. Finally, we provided a best evidence synthesis of the quality of the measurement properties of the instruments we identified, which makes for a distinct contribution to the integrated care literature.

Study Implications

The results of systematic reviews are generally useful for research, guideline development in evidence-based patient care, and policymaking.⁸⁰ Used with proper consideration of the review framework and its inherent

limitations, the results generated from our study have practical benefits for researchers, decision makers, and field practitioners alike.

First, the summary of the strength of evidence on the quality of the integrated care instruments makes key information on the instruments more accessible for end users interested in measuring integrated care. Many of these end users may have limited time and resources in gathering information and selecting an appropriate instrument from the pool of instruments available. Our results can also contribute to driving the development of guidelines for instrument selection. For now, the quality rating for the instruments in our review provides end users with an overview of the instruments available. In addition, the multilayered information provided on the integrated care constructs, instruments, and their measurement properties, are useful “virtual coordinates” that point to potential areas for further research. This baseline evidence on the measurement properties of integrated care instruments can guide future research on improving the quality of currently available instruments and can aid in investigating the less explored constructs and domains. Although our review was not designed to determine the best instrument for measuring integrated care, the instruments with strong evidence of positive findings for content and structural validity (online Appendix F) may be considered the prime candidates for an appropriate instrument for measuring the construct of integrated care.

Second, another insight from our results relates to our finding that instruments designed to measure clinical integration were mostly administered to patients. This is consistent with our framework showing clinical integration as an overarching dimension of person-focused care. On the other hand, it seems reasonable to measure professional integration among the health care providers. It is also interesting to note that a number of instruments measuring care integration (64%) and patient-centered care (58%) were administered to informal caregivers. This finding may be an indication of the potential role of informal caregivers in integrated care delivery.

Third, our work also draws attention to the need to improve not only the quality of instruments but also the reporting methods of primary studies on instrument development. Continued efforts to improve methodologies for systematic reviews of measurement properties would benefit subsequent research endeavors in this field. As new ideas and efforts to achieve a unified framework in the study of integrated care emerge, future studies on integrated care instruments may consider

conducting other systematic reviews. For example, performing a systematic review of the construct validity of integrated care instruments can provide a more focused assessment of the available instruments.

Finally, relevant investigations in the pipeline can potentially benefit from a systematic review of the theoretical frameworks underlying integrated care. A recent study that analyzed existing theoretical frameworks for examining care coordination¹⁰⁸ emphasized the importance of adopting theoretical frameworks in the development, implementation, and evaluation of care coordination. The use of theoretical frameworks ensures the congruence between the chosen interventions and the problems they intend to address. A systematic review of theoretical frameworks for studying integrated care can therefore be useful in selecting the appropriate framework for specific research priorities. Such a review may also help determine important considerations that would warrant the use of multiple frameworks.

Conclusion

The findings of our systematic review suggest that the quality of the measurement properties of instruments measuring integrated care may be improved with a thorough assessment of their measurement properties. The less-studied constructs (eg, continuity/comprehensive care, care coordination/case management) and domains (eg, professional, organizational, system, functional, normative integration) should be investigated further. Overall, the review findings can serve as a useful guide in the development and validation of instruments measuring integrated care.

The growing interest in integrated care guarantees more research in this field. Such dynamics properly put our work in perspective because despite the scale of our review, what we managed to capture and present is but a snapshot of the highly evolving landscape of integrated care measurement. The main results of our review suggest similar gaps in integrated care measurement from other reviews that underscore the lack of a standard measurement instrument and a unified framework for its measurement. While it seems logical to aim for an instrument that captures all the relevant dimensions of integrated care, some have argued whether one comprehensive instrument would indeed provide a better measure. In any case, it must be emphasized that the objectives of an integrated care assessment should be the main consideration in

selecting appropriate measures. This systematic review of the measurement properties of instruments does not directly answer how services can be integrated or how patients want care to be delivered. It also does not provide evidence of the effectiveness of integrated care (outcomes). Nevertheless, our systematic review provides a good starting point for steering quality metrics development in integrated care.

References

1. Divo MJ, Martinez CH, Mannino DM. Ageing and the epidemiology of multimorbidity. *Eur Respir J*. 2014;44(4):1055-1068.
2. Ezeh AC, Bongaarts J, Mberu B. Global population trends and policy options. *Lancet*. 2012;380(9837):142-148.
3. United Nations Population Fund (UNFPA), HelpAge International. *Ageing in the Twenty-First Century: A Celebration and a Challenge*. New York, NY, and London, England: UNFPA and HelpAge International; 2012.
4. Global Burden of Disease Study 2013 Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet*. 2015;386(9995):743-800.
5. Wagner EH, Davis C, Schaefer J, Von Korff M, Austin B. A survey of leading chronic disease management programs: are they consistent with the literature? *Manag Care Q*. 1999;7(3):56-66.
6. World Health Organization, Health Care for Chronic Conditions Team. *Innovative Care for Chronic Conditions: Building Blocks for Action*. Geneva, Switzerland: World Health Organization; 2002.
7. Leutz WN. Five laws for integrating medical and social services: lessons from the United States and the United Kingdom. *Milbank Q*. 1999;77(1):77-110, iv-v.
8. Kodner DL, Spreeuwenberg C. Integrated care: meaning, logic, applications, and implications—a discussion paper. *Int J Integr Care*. 2002;2:e12.
9. Leutz W. Reflections on integrating medical and social care: five laws revisited. *J Integr Care*. 2005;13(5):3-12.
10. Berwick DM, Nolan TW, Whittington J. The triple aim: care, health, and cost. *Health Aff (Millwood)*. 2008;27(3):759-769.
11. Kodner DL. All together now: a conceptual exploration of integrated care. *Healthc Q*. 2009;13(spec issue):6-15.

12. Shaw S, Rosen R, Rumbold B. *An Overview of Integrated Care in the NHS: What Is Integrated Care?* London, England: Nuffield Trust; 2011.
13. Hardy B, Mur-Veemanu I, Steenbergen M, Wistow G. Inter-agency services in England and the Netherlands; a comparative study of integrated care development and delivery. *Health Policy*. 1999;48(2):87-105.
14. Grone O, Garcia-Barbero M; WHO European Office for Integrated Health Care Services. Integrated care: a position paper of the WHO European Office for Integrated Health Care Services. *Int J Integr Care*. 2001;1:e21.
15. Burns LR, Pauly MV. Integrated delivery networks: a detour on the road to integrated health care? *Health Aff (Millwood)*. 2002;21(4):128-143.
16. Delnoij D, Klazinga N, Glasgow IK. Integrated care in an international perspective. *Int J Integr Care*. 2002;2:e04.
17. Ahgren B, Axelsson R. Evaluating integrated health care: a model for measurement. *Int J Integr Care*. 2005;5:e01; discussion e03, e09.
18. Fulop N, Mowlem A, Edwards N. *Building Integrated Care: Lessons From the UK and Elsewhere*. London, England: NHS Confederation; 2005.
19. Axelsson R, Axelsson SB. Integration and collaboration in public health—a conceptual framework. *Int J Health Plann Manage*. 2006;21(1):75-88.
20. Browne G, Kingston D, Grdisa V, Markle-Reid M. Conceptualization and measurement of integrated human service networks for evaluation. *Int J Integr Care*. 2007;7:e51.
21. Howarth ML, Haigh C. The myth of patient centrality in integrated care: the case of back pain services. *Int J Integr Care*. 2007;7:e27.
22. Edgren L. The meaning of integrated care: a systems approach. *Int J Integr Care*. 2008;8:e68.
23. Thomas P, Meads G, Moustafa A, Nazareth I, Stange KC, Donnelly Hess G. Combined horizontal and vertical integration of care: a goal of practice-based commissioning. *Qual Prim Care*. 2008;16(6):425-432.
24. Armitage GD, Suter E, Oelke ND, Adair CE. Health systems integration: state of the evidence. *Int J Integr Care*. 2009;9:e82.
25. Goodwin N. Integrated care as a scientific discipline: the need for more theory and new analytical methods. *Int J Integr Care*. 2010;10:e76.

26. Grdisa V. Conceptualizing and Measuring Integration: Tools and Resources for System Leaders [Powerpoint presentation]. Presented at: 2010 National Healthcare Leadership Conference; June 7-8, 2010; Winnipeg, Canada. <http://www.nhlc-cnls.ca/assets/grdisa%20presentation.pdf>. Accessed July 28, 2016.
27. Ahgren B, Axelsson R. A decade of integration and collaboration: the development of integrated health care in Sweden 2000-2010. *Int J Integr Care*. 2011;11(spec issue):e007.
28. Jiwani I, Fleury MJ. Divergent modes of integration: the Canadian way. *Int J Integr Care*. 2011;11(spec issue):e018.
29. Singer SJ, Burgers J, Friedberg M, Rosenthal MB, Leape L, Schneider E. Defining and measuring integrated patient care: promoting the next frontier in health care delivery. *Med Care Res Rev*. 2011;68(1):112-127.
30. Ahgren B. The art of integrating care: theories revisited. *Open Public Health J*. 2012;5:36-39.
31. Minkman MMN. The current state of integrated care: an overview. *J Integr Care*. 2012;20(6):346-358.
32. Nolte E, Knai C, Hofmarcher M, et al. Overcoming fragmentation in health care: chronic care in Austria, Germany and the Netherlands. *Health Econ Policy Law*. 2012;7(1):125-146.
33. Goodwin N. Understanding integrated care: a complex process, a fundamental principle. *Int J Integr Care*. 2013;13:e011.
34. Cash-Gibson L, Rosenmoller M. Project INTEGRATE—a common methodological approach to understand integrated health care in Europe. *Int J Integr Care*. 2014;14:e035.
35. Rumbold B, Shaw S. Shaping the future of integrated care: what can we learn from history? *Int J Integr Care*. 2010;10(6).
36. Shortell SM, Casalino LP. Health care reform requires accountable care systems. *JAMA*. 2008;300(1):95-97.
37. Kluge H, World Health Organization Division of Health Systems and Public Health. *Roadmap for Developing A Framework for Action Towards Coordinated/Integrated Health Services Delivery in the WHO European Region: An Overview*. July 29, 2013. http://www.euro.who.int/__data/assets/pdf_file/0007/195829/Roadmap-for-developing-a-Framework-for-Action-towards-Coordinated-Integrated-Health-Services-Delivery-in-the-WHO-European-Region.pdf?ua=1. Accessed July 28, 2016.
38. World Health Organization. WHO framework on integrated people-centred health services. <http://www.who.int/servicedeliverysafety/areas/people-centred-care/en/>. Accessed July 28, 2016.

39. Kreindler SA, Larson BK, Wu FM, et al. Interpretations of integration in early accountable care organizations. *Milbank Q.* 2012;90(3):457-483.
40. Vrijhoef HJ, Berbee R, Wagner EH, Steuten LM. Quality of integrated chronic care measured by patient survey: identification, selection and application of most appropriate instruments. *Health Expect.* 2009;12(4):417-429.
41. Ham C, Curry N. *Integrated Care: What Is It? Does It Work? What Does It Mean for the NHS?* London, England: The King's Fund; 2011.
42. Valentijn PP, Boesveld IC, van der Klauw DM, et al. Towards a taxonomy for integrated care: a mixed-methods study. *Int J Integr Care.* 2015;15:e003.
43. Valentijn PP, Vrijhoef HJ, Ruwaard D, Boesveld I, Arends RY, Bruijnzeels MA. Towards an international taxonomy of integrated primary care: a Delphi consensus approach. *BMC Fam Pract.* 2015;16:64.
44. Billings J. What do we mean by integrated care? A European interpretation. *J Integr Care.* 2005;13(5):13-20.
45. Gillies RR, Shortell SM, Anderson DA, Mitchell JB, Morgan KL. Conceptualizing and measuring integration: findings from the health systems integration study. *Hosp Health Serv Adm.* 1993;38(4):467-489.
46. Burns LR, Muller RW. Hospital-physician collaboration: landscape of economic integration and impact on clinical integration. *Milbank Q.* 2008;86(3):375-434.
47. Fisher ES, Staiger DO, Bynum JP, Gottlieb DJ. Creating accountable care organizations: the extended hospital medical staff. *Health Aff (Millwood).* 2007;26(1):w44-57.
48. Rittenhouse DR, Shortell SM, Fisher ES. Primary care and accountable care—two essential elements of delivery-system reform. *N Engl J Med.* 2009;361(24):2301-2303.
49. Kreindler SA, Dowd DA, Dana Star N, Gottschalk T. Silos and social identity: the social identity approach as a framework for understanding and overcoming divisions in health care. *Milbank Q.* 2012;90(2):347-374.
50. Curry N, Ham C. *Clinical and Service Integration: The route to Improved Outcomes.* London, England: The King's Fund; 2010. <https://www.kingsfund.org.uk/sites/files/kf/Clinical-and-service-integration-Natasha-Curry-Chris-Ham-22-November-2010.pdf>. Accessed July 28, 2016.
51. Lloyd J, Wait S. *Integrated Care: A Guide for Policymakers.* London, England: Alliance for Health and the Future; 2004.

52. van der Klauw D, Molema H, Grooten L, Vrijhoef H. Identification of mechanisms enabling integrated care for patients with chronic diseases: a literature review. *Int J Integr Care*. 2014;14:e024.
53. Valentijn PP, Schepman SM, Opheij W, Bruijnzeels MA. Understanding integrated care: a comprehensive conceptual framework based on the integrative functions of primary care. *Int J Integr Care*. 2013;13:e010.
54. de Silva D. *Helping Measure Person-Centred Care: A Review of Evidence About Commonly Used Approaches and Tools Used to Help Measure Person-Centred Care*. London, England: The Health Foundation; 2014.
55. Lawrence M, Kinn S. Defining and measuring patient-centred care: an example from a mixed-methods systematic review of the stroke literature. *Health Expect*. 2012;15(3):295-326.
56. Haggerty JL, Reid RJ, Freeman GK, Starfield BH, Adair CE, McKendry R. Continuity of care: a multidisciplinary review. *BMJ*. 2003;327(7425):1219-1221.
57. Rogers J, Curtis P. The concept and measurement of continuity in primary care. *Am J Public Health*. 1980;70(2):122-127.
58. Schultz EM, Pineda N, Lonhart J, Davies SM, McDonald KM. A systematic review of the care coordination measurement landscape. *BMC Health Serv Res*. 2013;13:119.
59. Schang L, Waibel S, Thomson S. *Measuring Care Coordination: Health System and Patient Perspectives*. London, England: Main Association of Austrian Social Security Institutions; 2013.
60. Doessing A, Burau V. Care coordination of multimorbidity: a scoping study. *J Comorbidity*. 2015;5(1):15-28.
61. Conklin A, Nolte E, Vrijhoef H. Approaches to chronic disease management evaluation in use in Europe: a review of current methods and performance measures. *Int J Technol Assess Health Care*. 2013;29(1):61-70.
62. Sternberg SB, Co JP, Homer CJ. Review of quality measures of the most integrated health care settings for children and the need for improved measures: recommendations for initial core measurement set for CHIPRA. *Acad Pediatr*. 2011;11(3 Suppl):S49-S58.e3.
63. Lyngso AM, Godtfredsen NS, Host D, Frolich A. Instruments to assess integrated care: a systematic review. *Int J Integr Care*. 2014;14:e027.
64. Strandberg-Larsen M, Krasnik A. Measurement of integrated healthcare delivery: a systematic review of methods and future research directions. *Int J Integr Care*. 2009;9:e01.

65. Valentine MA, Nembhard IM, Edmondson AC. Measuring teamwork in health care settings: a review of survey instruments. *Med Care*. 2015;53(4):e16-e30.
66. Graham C, Killpack C, Raleigh V, Redding D, Thorlby R, Walsh J. *Options Appraisal on the Measurement of People's Experiences of Integrated Care*. Oxford, England: The King's Fund, National Voices, the Nuffield Trust, and Picker Institute Europe; 2013.
67. King J, Gibbons E, Graham C, Walsh J. *Developing Measures of People's Self-Reported Experiences of Integrated Care*. Oxford, England: Picker Institute Europe, University of Oxford; 2013.
68. Minkman MMN. Developing integrated care: towards a development model for integrated care. *Int J Integr Care*. 2012;(12): e97.
69. Minkman MMN, Vermeulen RP, Ahaus KTB, Huijsman R. The implementation of integrated Care: the empirical validation of the Development Model for Integrated Care. *BMC Health Serv Res*. 2011;11:177.
70. Uijen AA, Heinst CW, Schellevis FG, et al. Measurement properties of questionnaires measuring continuity of care: a systematic review. *PLoS One*. 2012;7(7):e42256.
71. Berg M. Measuring the value of healthcare delivery: a call to action. BMJ Outcomes website. <http://outcomes.bmj.com/index.php/journal/measuring-the-value-of-healthcare-delivery-a-call-to-action>. Published April 2015. Accessed July 28, 2016.
72. Donabedian A. Evaluating the quality of medical care, 1966. *Milbank Q*. 2005;83(4):691-729.
73. Taylor EF, Lake T, Nysenbaum J, Peterson G, Meyers D. *Coordinating Care in the Medical Neighborhood: Critical Components and Available Mechanisms* (White Paper). Washington, DC: US Department of Health and Human Services, Agency for Healthcare Research and Quality; 2011.
74. Goodwin N, Smith J, Davies A, et al. *Integrated Care for Patients and Populations: Improving Outcomes by Working Together*. London, England: The King's Fund; 2012.
75. Springer JF, Phillips J. *The Institute of Medicine Framework and Its Implication for the Advancement of Prevention Policy, Programs and Practice*. Santa Rosa, CA: Center for Applied Research Solutions; 2007.
76. Hofmarcher MM, Oxley H, Rusticelli E. *Improved Health System Performance Through Better Care Coordination*. Paris, France: OECD Health Working Papers; 2007.

77. Strandberg-Larsen M. Measuring integrated care. *Dan Med Bull.* 2011;58(2):B4245.
78. Psychometric properties [definition]. *Stedman's Medical Dictionary for the Health Professions and Nursing*. 7th ed. Philadelphia, PA: Lippincot Williams & Wilkins; 2012.
79. VU University Medical Center. Systematic reviews of measurement properties. COSMIN website. [http://www.cosmin.nl/Systematic reviews of measurement properties.html](http://www.cosmin.nl/Systematic%20reviews%20of%20measurement%20properties.html). Accessed July 28, 2016.
80. de Vet HC, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine: A Practical Guide*. Cambridge, England: Cambridge University Press; 2011.
81. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60(1):34-42.
82. Mokkink L, Terwee C, Stratford P, et al. Evaluation of the methodological quality of systematic reviews of health status measurement instruments. *Qual Life Res.* 2009;18(3):313-333.
83. 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):1115-1123.
84. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res.* 2012;21(4):651-657.
85. Shultz KS, Whitney DJ. *Measurement Theory in Action: Case Studies and Exercises*. Thousand Oaks, CA: SAGE Publications Inc.; 2005.
86. 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.
87. Mokkink LB, Terwee CB, Patrick 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):539-549.
88. Mokkink LB, Terwee CB, Knol DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. *BMC Med Res Methodol.* 2010;10:22.
89. Mokkink LB, Terwee CB, Patrick DL, et al. *COSMIN Checklist Manual*. Amsterdam and Leiden, the Netherlands: EMGO

- Institute for Health and Care Research, VU University Medical Center, Anna Foundation; 2012.
90. Welcome at the Knowledgecenter Measurement Instruments (KMIN) of the VU University Medical Center. KMIN website. <http://www.kmin-vumc.nl/index.php?change=1&lang=en>. Accessed July 28, 2016.
 91. Triemstra M, Winters S, Kool RB, Wieggers TA. Measuring client experiences in long-term care in The Netherlands: a pilot study with the Consumer Quality Index Long-term Care. *BMC Health Serv Res.* 2010;10:95.
 92. Stubbe JH, Gelsema T, Delnoij DM. The Consumer Quality Index Hip Knee Questionnaire measuring patients' experiences with quality of care after a total hip or knee arthroplasty. *BMC Health Serv Res.* 2007;7:60.
 93. Stubbe JH, Brouwer W, Delnoij DM. Patients' experiences with quality of hospital care: the Consumer Quality Index Cataract Questionnaire. *BMC Ophthalmol.* 2007;7:14.
 94. Howie JG, Heaney DJ, Maxwell M, Walker JJ, Freeman GK. Developing a 'consultation quality index' (CQI) for use in general practice. *Fam Pract.* 2000;17(6):455-461.
 95. Damman OC, Hendriks M, Sixma HJ. Towards more patient centred healthcare: A new Consumer Quality Index instrument to assess patients' experiences with breast care. *Eur J Cancer.* 2009;45(9):1569-1577.
 96. Booi J, Zegers M, Evers PM, Hendriks M, Delnoij DM, Rademakers JJ. Improving cancer patient care: development of a generic cancer Consumer Quality Index questionnaire for cancer patients. *BMC Cancer.* 2013;13:203.
 97. Berendsen AJ, Groenier KH, de Jong GM, et al. Assessment of patient's experiences across the interface between primary and secondary care: Consumer Quality Index continuum of care. *Patient Educ Couns.* 2009;77(1):123-127.
 98. Haggerty JL, Beaulieu MD, Pineault R, et al. Comprehensiveness of care from the patient perspective: comparison of primary healthcare evaluation instruments. *Healthc Policy.* 2011;7(spec issue):154-166.
 99. Flocke SA. Measuring attributes of primary care: development of a new instrument. *J Fam Pract.* 1997;45(1):64-74.
 100. Haggerty JL, Levesque JF, Santor DA, et al. Accessibility from the patient perspective: comparison of primary healthcare evaluation instruments. *Healthc Policy.* 2011;7(spec issue):94-107.
 101. Haggerty JL, Burge F, Beaulieu MD, 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.
102. Beaulieu MD, Haggerty JL, Beaulieu C, et al. Interpersonal communication from the patient perspective: comparison of primary healthcare evaluation instruments. *Healthc Policy*. 2011;7(spec issue):108-123.
 103. Institute of Medicine Committee on Psychosocial Services to Cancer Patients/Families in a Community Setting. The National Academies Collection: Reports funded by National Institutes of Health. In: Adler NE, Page AEK, eds. *Cancer Care for the Whole Patient: Meeting Psychosocial Health Needs*. Washington, DC: National Academies Press, National Academy of Sciences; 2008.
 104. Understanding the differences between Medline and Embase. McGill University Health Centre website. <http://www.muhclibraries.ca/understanding-the-differences-between-medline-and-embase>. Accessed July 28, 2016.
 105. Woods D, Trewheellar K. Medline and Embase complement each other in literature searches. *BMJ*. 1998;316(7138):1166.
 106. Morrison A, Polisen J, Husereau D, et al. The effect of English-language restriction on systematic review-based meta-analyses: a systematic review of empirical studies. *Int J Technol Assess Health Care*. 2012;28(2):138-144.
 107. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. 2010;63(7):737-745.
 108. Van Houdt S, Heyrman J, Vanhaecht K, Sermeus W, De Lepeleire J. An in-depth analysis of theoretical frameworks for the study of care coordination. *Int J Integr Care*. 2013;13:e024.

Funding/Support: This study is part of a project funded by the Singapore Ministry of Education (MOE) Academic Research Fund (AcRF) Tier 1 Grant awarded to HJMV and LYW.

Conflict of Interest Disclosures: All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. No disclosures were reported.

Acknowledgments: We thank all the authors of the primary studies included in this systematic review, especially those who shared their full-text articles, which were otherwise unavailable.

Address correspondence to: Hubertus J.M. Vrijhoef, MSc, PhD, Department of Patient & Care, Maastricht University Medical Center, PO Box 5800, 6202 AZ Maastricht, the Netherlands (e-mail: b.vrijhoef@mumc.nl).

Supplementary Material

Additional supporting information may be found in the online version of this article at [http://onlinelibrary.wiley.com/journal/10.1111/\(ISSN\)1468-0009](http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1468-0009):

Notes on Appendices

Appendix A. Final Search Strategy

Appendix B. Key Characteristics of Articles Included in the Systematic Review

Appendix C. Key Characteristics of Instruments Identified

Appendix D. Methodological Quality of Studies per Measurement Property

Appendix E. Overall Quality (Adequacy) of Measurement Properties per Index Instrument

Appendix F. Instruments With Evidence of Adequate Measurement Properties

Appendix G. Instruments Grouped According to Domain