

RESEARCH ARTICLE

Open Access



Home care quality indicators based on the Resident Assessment Instrument-Home Care (RAI-HC): a systematic review

Aylin Wagner^{1*†}, René Schaffert^{1†}, Nathalie Möckli^{2†}, Franziska Zúñiga^{2†} and Julia Dratva^{1,3†}

Abstract

Background: One way of measuring the quality of home care are quality indicators (QIs) derived from data collected with the Resident Assessment Instrument-Home Care (RAI-HC). In order to produce meaningful results for quality improvement and quality comparisons across home care organizations (HCOs) and over time, RAI-HC QIs must be valid and reliable. The aim of this systematic review was to identify currently existing RAI-HC QIs and to summarize the scientific knowledge on the validity and reliability of these QIs.

Methods: A systematic review was performed using the electronic databases PubMed, CINAHL, Embase, PsycINFO and Cochrane Library. Studies describing the development process or the psychometric characteristics of RAI-HC QIs were eligible. The data extraction involved a general description of the included studies as well as the identified RAI-HC QIs and information on validity and reliability. The methodological quality of the identified RAI-HC QI sets was assessed using the Appraisal of Indicators through Research and Evaluation (AIRE) instrument.

Results: Four studies out of 659 initial hits met the inclusion criteria. The included studies described the development and validation process of three RAI-HC QI sets comprising 48 unique RAI-HC QIs, which predominantly refer to outcome of care. Overall, the validity and reliability of the identified RAI-HC QIs were not sufficiently tested. The methodological quality of the three identified RAI-HC QI sets varied across the four AIRE instrument domains. None of the QI sets reached high methodological quality, defined as scores of 50% and higher in all four AIRE instrument domains.

Conclusions: This is the first review that systematically summarized and appraised the available scientific evidence on the validity and reliability of RAI-HC QIs. It identified insufficient reporting of RAI-HC QIs validation processes and reliability as well as missing state-of-the-art methodologies. The review provides guidance as to what additional validity and reliability testing are needed to strengthen the scientific soundness of RAI-HC QIs. Considering that RAI-HC QIs are already implemented and used to measure and compare quality of home care, further investigations on RAI-HC QIs reliability and validity is recommended.

Keywords: Systematic review, Quality indicators, Home care, Home-based care, Nursing, InterRAI, Validity, Reliability

* Correspondence: aylin.wagner@zhaw.ch

Aylin Wagner, René Schaffert, Nathalie Möckli, Franziska Zúñiga, Julia Dratva (shared last authorship)

¹ZHAW Zurich University of Applied Sciences, School of Health Professions, Institute of Health Sciences, Technikumstrasse 71, 8401 Winterthur, Switzerland

Full list of author information is available at the end of the article



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

The change of populations' age-structure has a significant impact on health systems worldwide and, in particular, poses challenges for home care [1]. Home care in the context of this study is defined as medical and personal care provided by professional nursing staff within clients' own homes. Globally, the number of people aged 60 and older is expected to double by 2050 [2]. As larger demographic cohorts enter old age and life expectancy increases, more people will live with chronic illnesses, multi-morbidity, as well as functional and cognitive impairments [3]. Findings have shown that the large majority of older people in need of care prefer to remain in their known physical and social environment for as long as possible, leading to increased demand for home care [4]. In order to satisfy peoples' preferences and to reduce costs of long-term institutional care, many countries have promoted home care in recent decades by shifting resources accordingly [1]. Given the growing importance of home care, it is essential to assess and monitor the quality of the delivered care.

Quality indicators (QIs) are increasingly used to measure, monitor and evaluate health care quality by assessing particular structures, processes, or outcomes. QIs can point to areas where the quality of care is suboptimal, subsequently allowing priorities to be set for quality improvement [5]. Moreover, QIs are used for comparisons of health care quality and thus enable national or international benchmarking [6]. Similar to other measuring instrument, QIs should meet certain quality criteria such as relevance and feasibility and be evaluated for their scientific strength, i.e. their validity and reliability [7]. QIs that do not meet these quality criteria can result in inadequate and misleading information.

In many countries in North America, Europe and Asia-Pacific, quality of home care is assessed with QIs based on the Resident Assessment Instrument-Home Care (RAI-HC or interRAI-HC). RAI-HC was developed in 1994 by the multinational research consortium interRAI. The instrument is a standardized assessment tool to assess long-stay home care clients' health status, need for care, and basic background information on housing and informal caregivers. The instrument RAI-HC [8–13] and its clinically based scales (e.g. Cognitive Performance Scale, Depression Rating Scale) [14–20] have been validated in several international studies. Although the principal intended use of RAI-HC is to plan care provision, RAI-HC items and scales are also used to derive process and outcome QIs [21]. These RAI-HC QIs are rate-based indicators, i.e. defined by a numerator and denominator, and measure processes or outcomes expected to occur with a certain frequency [22].

To date, no systematic review has been undertaken to summarize the scientific soundness, such as the validity

and reliability of RAI-HC QIs, despite the fact that these indicators are implemented in several countries and applied by researchers to measure and compare the quality of home care [23–26]. The only previous overview and quality assessment of QIs for community care by Joling et al. [27] included RAI-HC QIs only partially, as it focused on QIs specifically developed for older people or applied in an older aged sample (i.e. 65 years or older). Therefore, this systematic review aimed to (i) identify all current existing RAI-HC QIs and to (ii) summarize the scientific evidence of RAI-HC QIs validity and reliability.

Methods

The systematic review was conducted in compliance with The Cochrane Handbook for Systematic Reviews of Interventions [28]. The protocol for the systematic review has been published on PROSPERO (2018: CRD42018110948) and is available at https://www.crd.york.ac.uk/prospERO/display_record.php?RecordID=110948.

Search strategy

The search was carried out using five electronic databases: PubMed, CINAHL, Embase, PsycINFO and Cochrane Library on June 26, 2018. An update of the search was conducted on August 20, 2019, resulting in no additional eligible articles. The search strategy involved both keywords and the Medical Subject Headings (MeSH) combined with the appropriate Boolean connectors (see Additional file 1). In addition, the reference lists of the eligible studies were manually searched for additional relevant articles that had not been identified in the electronic database. Furthermore, we searched for grey literature on websites of relevant organizations (e.g. www.interrai.org) and contacted the study authors of two included articles, namely Burla et al. [29] and Morris et al. [30], to obtain additional information on the QIs definitions used in their study.

Inclusion and exclusion criteria

Studies were included if they fulfilled the following criteria: (i) the study was conducted in the home care setting, (ii) included adults aged 18 years and older, and (iii) described the development process of RAI-HC QIs or evaluated the psychometric characteristics of RAI-HC QIs.

Studies were excluded if they (i) used RAI-HC or its scales without explicitly using the RAI-HC QIs, (ii) applied already developed RAI-HC QIs for quality measure, (iii) validated the RAI-HC or its scales but not RAI-HC QIs, and (iv) focused on specialized home care for mental health or palliative care. We excluded studies focusing on mental health and palliative care because the needs of these home care clients are different from those

of general long-stay home care clients. Therefore, specialized assessment instruments such as the interRAI Palliative Care (PC) and the interRAI Community Mental Health (CMH) are available to assess the needs of mental health and palliative care clients and to plan their care provision. Only recently, QIs have been developed to measure the quality in these contexts [31, 32].

Screening

The studies identified by the electronic search were entered into the reference management software Endnote X8, and duplicates were removed. The articles were independently screened by two authors (AW and FZ) according to the inclusion criteria first by title and abstract. Non-eligible studies were removed at this stage. Publications included after the title/abstract screening underwent concurrent full-text screening by two reviewers (AW and FZ) for definitive inclusion. Any disagreements that arose between the reviewers were resolved through discussion until consensus was reached.

Data extraction

Two data extraction forms were developed. First, a structured form was used to describe the included studies with respect to relevant information regarding RAI-HC QIs. The following data were extracted: first author, year of publication, country, study aim, study population, sample size, name of QI set, the number of QIs in the set and a short description of the development and validation process of the QI set. Second, a structured form was used to extract and summarize the identified RAI-HC QIs. For each QI, the following data were extracted: QI description such as name, type (prevalence or incidence), name of the corresponding QI set, and data regarding validity and reliability. Furthermore, the QIs were classified by the study authors (AW and FZ) according to their measure level (outcome or process). Missing data regarding QI definitions were requested from study authors.

Methodological assessment

We used the Appraisal of Indicators through Research and Evaluation (AIRE) instrument for the methodological assessment of the RAI-HC QIs identified in the articles [33]. AIRE is a validated instrument for a critical appraisal of QIs and has been used in previous scientific publications on QIs [27, 34–36]. The AIRE instrument comprises 20 items, subdivided into four domains: 1. Purpose, relevance and organizational context, 2. Stakeholder involvement, 3. Scientific evidence, and 4. Additional evidence, formulation and usage. Two authors (AW and NM) independently appraised all included QI sets with the AIRE instrument. Information from the

included articles as well as QI definitions received from study authors on request were used for the assessment. Standardized scores per domain were calculated. Scores range between 0 and 100%, with a higher score indicating a higher methodological level. A high methodological quality is defined by a score of 50% or higher, which correlates with an overall “agree” or “strongly agree” for each domain.

Results

Search results

The systematic review identified 659 potentially relevant studies. The PRISMA flow diagram for the study selection process and reasons for exclusion are shown in Fig. 1. After removal of duplicates and title/abstract screening, 21 studies were potentially relevant. Four studies met the selection criteria after full-text screening [29, 30, 37, 38]. Reference tracking of the eligible studies identified no additional article. The four included studies describe three RAI-HC QI sets comprising 48 unique RAI-HC QIs.

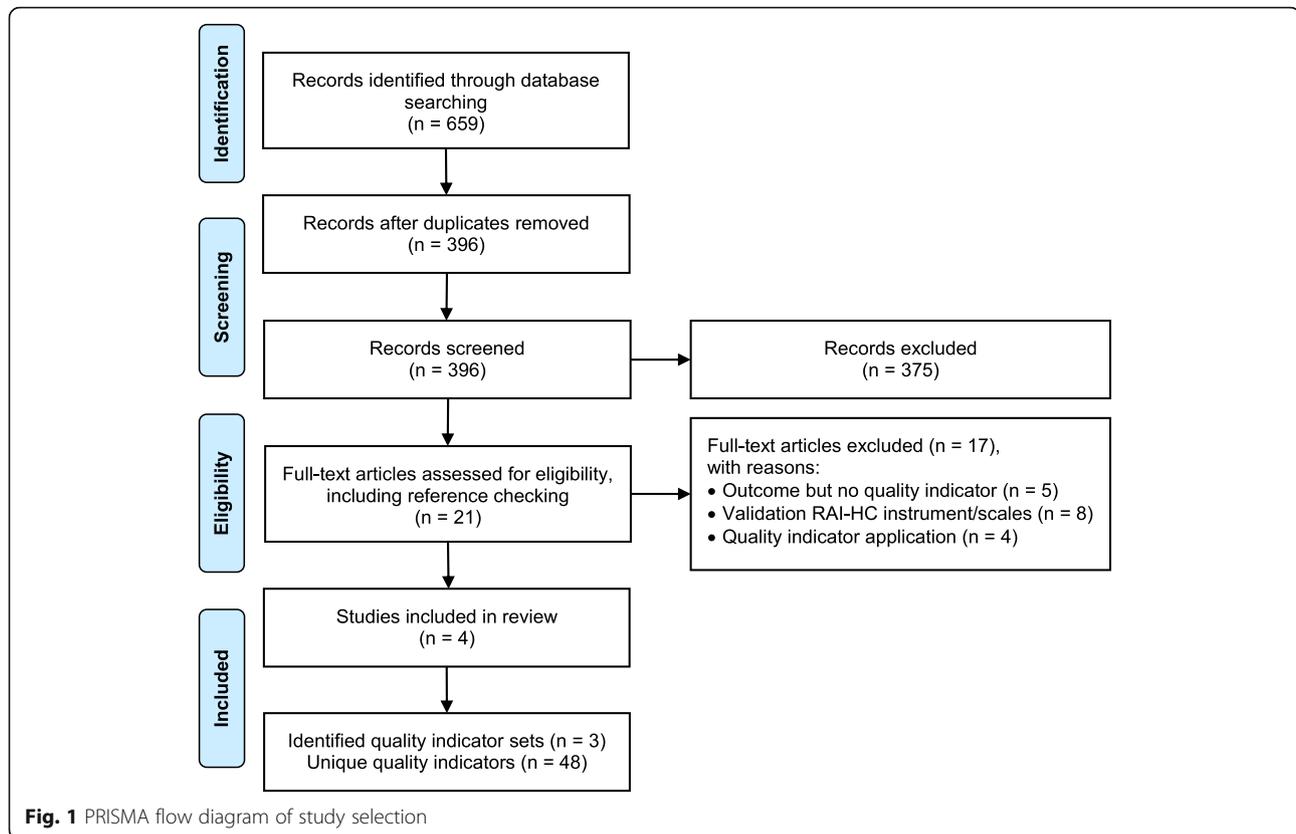
Description of studies

Table 1 shows the main characteristics of the selected papers. Three studies specified the development and validation process of separate RAI-HC QI sets, namely, interRAI's 1st generation QI set developed by Hirdes et al. [37], the Swiss RAI-HC QI set developed by Burla et al. [29], and interRAI's 2nd generation QI set developed by Morris et al. [30]. One study by Dalby et al. [38] explored the effects of risk adjustment for interRAI's 1st generation QI set. The included studies were not primarily aimed at examining the validity and reliability of RAI-HC QIs, and only partially reported such results.

Development and validation process of quality indicator sets

InterRAI's 1st generation quality indicator set

Hirdes et al. [37] described the development and validation process of interRAI's 1st generation QI set. Initially, a literature review was conducted to identify candidate QIs used in other care settings. In addition, expert meetings and focus groups with health professionals and older adults were conducted to identify further candidate QIs. The literature review (e.g. search string, results, synthesis) and process of the focus groups were not described in the article. In total, 74 candidate QIs were generated which were first prioritized by the investigators according to their relevance to different types of home care clients and second, ranked in terms of their appropriateness for the different types of home care clients. The average ranks were then used to reduce the number of candidate QIs. The article offers no information on the ranking process, the average ranks and



the criteria determining when a QI was considered inappropriate. The QIs were empirically tested with regard to relative frequencies, variation and denominator size based on data from 14,293 home care clients from Ontario and Michigan. QIs with a relative frequency of less than 5% and more than 95%, respectively, and too little variation (interquartile range) among HCOs were excluded. The article provides no information on the basis of which criteria QI variation was considered as insufficient. Based on the overall study results, a final set of 22 QIs was defined, for which risk adjustment methods such as client-level covariates and an agency-level adjuster, namely, the Agency Intake Profile (AIP), has been quantitatively evaluated [37].

Dalby et al. [38] further explored the effects of risk adjustment for interRAI's 1st generation QI set. Based on data of 22 HCOs in Ontario and the Winnipeg Regional Health Authority (WRHA) in Manitoba, three types of risk adjustment methods were applied, namely, client covariates only, client covariates plus AIP, and client covariates plus the intake Case Mix Index associated with the Resource Utilization Groups version III for Home Care methodology [39]. Based on the three approaches, risk adjustment showed substantial effects on the organization level but only small effects on the regional level. On the regional level, the risk adjustment process

minimized the differences in QI rates between Ontario and the WRHA compared with the unadjusted rates. On the organization level, risk adjustment had an impact on agency rankings across the set of QIs. While the HCOs in Ontario benefited from the risk adjustment, i.e. they were less likely to be ranked among the worst performers, the opposite was true for the HCOs in the WRHA [38].

The Swiss RAI-HC quality indicator set

The development and validation process of the Swiss RAI-HC QI set was described by Burla et al. [29]. Based on interRAI's 1st generation QI set and by creating new QIs for the Swiss context with support of various experts, 29 candidate QIs were chosen. The QIs were rated according to their appropriateness of measuring home care quality in focus groups with health care professionals from HCOs using the nominal group technique (NGT). The rating process and the QI ratings were presented in the paper. Relative frequencies of all candidate QIs and the variation of 24 QIs (due to small sample size) were examined based on data from 1808 home care clients from Switzerland. QIs with a relative frequency of less than 5% or more than 95% and/or a low variation (interquartile difference less than 6%) were specified as inadequate. Furthermore, due to small sample size

Table 1 Included articles in review

1st author, Year of publication	Country	Study aim	Study population (sample size)	Name of QI set	Number of QIs	Development and validation process of QI set
Hirdes, 2004 [37]	Canada and USA	To develop home care QIs based on RAI-HC.	Persons receiving home care services for at least 30 days ($n = 14,293$ clients)	interRAI's 1st generation QI set	Total: 22 Outcome: 19 Process: 3	Step 1: Identification of candidate QIs based on literature review, focus groups and expert meetings. Step 2: Ranking of QIs by investigators. Step 3: Empirical testing of QIs (denominator size, relative frequency and variation) and development of risk adjustment based on cross-national data.
Dalby, 2005 [38]	Canada	To explore the effect of risk adjustment for interRAI's 1st generation QI set.	Persons aged 18 years and older receiving home care services ($n = 22$ HCOs)	interRAI's 1st generation QI set	Total: 22 Outcome: 19 Process: 3	No QI development, but investigation of three approaches of risk adjustment and comparison of unadjusted and risk-adjusted QI rates.
Burla, 2010 [29]	Switzerland	To develop and examine home care QIs based on RAI-HC for Switzerland.	Persons aged 18 years and older receiving home care services ($n = 1808$ clients, 45 HCOs)	Swiss RAI-HC QI set	Total: 29 Outcome: 27 Process: 2	Step 1: Identification of candidate QIs based on interRAI's 1st generation QI set and development of new QIs for the Swiss context. Step 2: Rating of QIs in focus groups with health care professionals (nominal group technique). Step 3: Empirical testing of QIs (denominator size, relative frequency and variation) based on Swiss home care data. Step 4: Interrater reliability testing of subset of QIs.
Morris, 2013 [30]	Europe, Canada, USA	To develop 2nd generation home care QIs based on RAI-HC.	Persons aged 65 years and older receiving home care services ($n = 335,544$ clients, 1654 HCOs)	interRAI's 2nd generation QI set	Total: 23 Outcome: 22 Process: 1	Step 1: Identification of candidate QIs based on interRAI's 1st generation QI set and QIs from other settings. Step 2: Evaluation of QIs by representatives of HCOs in one-on-one discussions and focus groups. Review of QIs by interRAI's cross-national program development committee. Step 3: Empirical testing of QIs (denominator size, relative frequency, variation, factor analysis) and development of risk adjustment based on cross-national data.

Abbreviations: QI Quality indicator, RAI-HC Resident Assessment Instrument-Home Care, HCO Home care organization

interrater reliability was analyzed in only 18 QIs. For this purpose, 24 home care clients were independently assessed by two assessors. The results of the expert rating, frequencies, variation, and interrater reliability were summarized and only QIs that met at least three of the four criteria were defined as appropriate, resulting in a final core-set of 19 QIs [29].

InterRAI's 2nd generation quality indicator set

Morris et al. [30] described the development and validation process of interRAI's 2nd generation QI set. A list of 64 candidate QIs was compiled including both 1st generation QIs as well as newly designed QIs, drawn from lists of QIs that interRAI has considered for home care, post-acute care, and long-term care. The QIs were empirically tested with regard to relative frequencies and variation based on a sample of 335,544 home care clients from the U.S., Canada and Europe. QIs with a relative frequency of less than 3% were excluded. Variation was examined by comparing

scores of top performing HCOs (5th percentile) with scores of the poorest performing HCOs (95th percentile). QIs with less than a two-fold difference in scores from the 5th to 95th percentile, thus not discriminatory enough, were dropped. Additionally, factor analysis was conducted for eight functional QIs measuring improvement and decline in cognition, communication, activities of daily living (ADL) and instrumental activities of daily living (IADL) to confirm that functional decline and improvement QIs say something different about the performance of HCOs. The QIs were further evaluated by representatives of HCOs in focus groups and one-on-one discussions to determine if the QIs could be influenced by efforts of HCOs. The remaining QIs were reviewed regarding face validity by 16 members of interRAI's cross-national program development committee. Each QI had to be approved by at least 70% of the members to remain on the QI list. Rating results, rating criteria and the exact method of consensus are not indicated in the article. The definite QI set comprised

23 QIs. For the QI set, a new approach of risk adjustment was employed with more complex covariate structures, a longer list of covariates, and direct stratification [30].

Validity and reliability of RAI-HC quality indicators

Table 2 gives an overview of the characteristics of the 48 identified RAI-HC QIs and summarizes findings on their validity and reliability. Face validity was examined for all identified QIs based on expert opinion. However, only Burla et al. [29] described rating results for 29 Swiss RAI-HC QIs, with seven QIs rated as inappropriate to measure quality of home care. Reliability was addressed explicitly by Burla et al. [29] providing results of interrater reliability testing for the RAI-HC items used for the calculation of 18 Swiss RAI-HC QIs. Interrater reliability was calculated by Burla et al. [29] using Kappa and Yules. The higher Kappa/Yules (range 0–1), the higher the agreement between the two independent assessors. Twelve QIs showed moderate (Kappa/Yules values 0.40–0.59) or good interrater reliability (Kappa/Yules values ≥ 0.60) and six QIs had insufficient interrater reliability (Kappa/Yules values < 0.40) [29, 40].

Methodological characteristics of RAI-HC quality indicator sets

The methodological quality of the three identified RAI-HC QI sets varied according to the AIRE instrument domain scores (see Table 3). The AIRE instrument domain ratings ranged from 0 to 69%. None of the QI sets reached high methodological quality, defined as scores of 50% or higher in all four AIRE instrument domains.

InterRAI's 1st generation QI set [37] and the Swiss RAI-HC QI set [29] scored 50% or higher in the first AIRE instrument domain demonstrating good evidence for "purpose, relevance and organizational context". InterRAI's 1st generation QI set [37] and the Swiss RAI-HC QI set [29] scored poorly in the domain "Stakeholder involvement" due to a lack of involvement of relevant stakeholders at some stage of the development process. The three QI sets scored between 0 and 11% in the domain "Scientific evidence". None of the three studies performed a systematic review to investigate evidence-based guidelines supporting QIs nor did they examine the relationships between care processes and outcomes. The domain "Additional evidence, formulation and usage" indicated for interRAI's 1st and 2nd generation QI set [30, 37] a good methodological quality with scores of 50% or higher.

Discussion

In this systematic review, the three identified RAI-HC QI sets contained a total of 48 unique QIs covering various health domains and predominantly referring to outcome of care. To be able to draw valid conclusions from

QIs, it is relevant to establish the validity and reliability of the QIs. The methodological assessment of the three QI sets, however, indicated relatively low methodological quality and a lack of evidence of validity and reliability.

QI's validity, such as face, content, construct, and criterion aspects, should either be supported by scientific literature or be examined. When addressing scientific evidence, it is recommended to follow a systematic approach and to search both for scientific as well as grey literature, and not only to identify articles regarding the validity of QIs but also articles that discuss the outcome of interest [41]. However, this review showed that the above recommendations were not applied in the development processes of extant RAI-HC QIs. The majority of identified RAI-HC QIs were adopted from other care settings or other interRAI QI sets, and in none of the included studies was a systematic review carried out to identify candidate QIs or to address scientific evidence on QIs.

In practice, and in the absence of scientific literature on QIs, face validity of QIs is often assessed based on the opinions and experience of experts [42]. Commonly used structured consensus techniques for QI development combining expert opinion with available evidence are the Delphi techniques [43], the RAND/UCLA Appropriateness Method [44] and the NGT [45]. The advantages of these approaches are that experts can be included anonymously and the undue social influence processes toward expert consensus can be minimized [42]. In all three identified RAI-HC QI sets, experts were involved to evaluate the face validity of QIs, however, such a structured approach was only applied by Burla et al. [29] using the NGT in the development process of the Swiss RAI-HC QIs. Hirdes et al. [37] did not report on the exact process of assessing face validity for interRAI's 1st generation QIs and the consensus process used by Morris et al. [30] for interRAI's 2nd generation QIs was incompletely described. Considering that consensus for both interRAI QI sets was reached in an unstructured way and by face-to-face discussion only, rating results could be biased toward the opinions of influential or persuasive experts.

Other forms of validity such as content, construct and criterion validity of RAI-HC QIs have not been examined. Admittedly, similar to nursing home QIs [46, 47], criterion validity of home care QIs can be difficult to measure because few, if any, valid "gold standard" measurements exist that can be used for comparison. While some of the identified RAI-HC QIs are calculated using validated RAI-HC scales (e.g. Cognitive Performance Scale, Depression Rating Scale and Pain Scale) [14, 16–18], the use of validated scales does not necessarily ensure the validity of the corresponding QI. RAI-HC scales have not been developed to evaluate quality of

Table 2 Characteristics of identified RAI-HC quality indicators

QI	Measure level ^a	Type ^b	QI set affiliation ^c			Face validity ^d			Interrater reliability ^e
			interRAI 1st	Swiss RAI-HC	interRAI 2nd	interRAI 1st	Swiss RAI-HC	interRAI 2nd	
ADL	O	I	✓	✓		C	A		≥ 0.60
ADL decline	O	I			✓			B	
ADL improvement	O	I			✓			B	
IADL	O	I		✓			A		≥ 0.60
IADL decline	O	I			✓			B	
IADL improvement	O	I			✓			B	
Rehabilitation potential and no therapies	P	P	✓	(✓)		C	A		N/A
Decline independency	O	P		✓			A		N/A
Difficulties in communication	O	I	✓	✓		C	A		N/A
Communication decline	O	I			✓			B	
Communication improvement	O	I			✓			B	
Bladder incontinence	O	I	✓	✓		C	D		≥ 0.60
Bladder decline	O	I			✓			B	
Bladder improvement	O	I			✓			B	
Bowel incontinence	O	I		(✓)			D		≥ 0.60
Obstipation	O	I		(✓)			A		≥ 0.60
Skin ulcer	O	I	✓	✓		C	A		< 0.40
Mouth problems	O	P		✓			A		0.40–0.59
Cognitive function	O	I	✓	✓		C	D		0.40–0.59
Cognitive decline	O	I			✓			B	
Cognitive improvement	O	I			✓			B	
Delirium	O	P	✓	(✓)		C	A		N/A
Negative mood	O	P	✓	✓		C	A		< 0.40
Mood decline	O	I			✓			B	
Mood improvement	O	I			✓			B	
Negative mood without intervention	O	P		(✓)			A		< 0.40
Social isolation with distress	O	P	✓	✓	✓	C	A	B	0.40–0.59
Does not go out but used to	O	P			✓			B	
Continued caregiver distress	O	P		✓	✓		A	B	< 0.40
Falls	O	P	✓	✓	✓	C	D	B	≥ 0.60
Neglect or abuse	O	P	✓	✓		C	A		N/A
Injuries	O	P	✓		✓	C		B	
Hospital, ED, emergent care	O	P	✓		✓	C		B	
Daily severe pain	O	P	✓	✓	✓	C	A	B	0.40–0.59
Inadequate pain control	O	P	✓	✓	✓	C	A	B	< 0.40
Pain improvement	O	I			✓			B	
Dehydration	O	P	✓	✓		C	A		0.40–0.59
Weight loss	O	P	✓	✓	✓	C	A	B	N/A
Weight change (undesired)	O	I		(✓)			D		N/A
Weight change (unfavourable)	O	I		(✓)			D		N/A
Inadequate meals	O	P	✓			C			

Table 2 Characteristics of identified RAI-HC quality indicators (Continued)

QI	Measure level ^a	Type ^b	QI set affiliation ^c			Face validity ^d			Interrater reliability ^e
			interRAI 1st	Swiss RAI-HC	interRAI 2nd	interRAI 1st	Swiss RAI-HC	interRAI 2nd	
Difficulty in locomotion and no assistive device	O	P	✓			C			
Impaired locomotion in home health status at a single point in time	O	I	✓	✓		C	A	≥ 0.60	
No medication review by MD	P	P	✓	✓		C	A	N/A	
Inconsistent drug intake	O	P		(✓)			A	N/A	
Hearing	O	P		(✓)			A	< 0.40	
Eyesight	O	P		(✓)			D	N/A	
No flu vaccination	P	P	✓		✓	C		B	

Abbreviations: ADL Activities of daily living, ED Emergency department, IADL Instrumental activities of daily living, MD Medical doctor, QI Quality indicator, RAI-HC Resident Assessment Instrument-Home Care

^a O = Outcome, P = Process, classified by authors

^b I = Incidence-based measure (measures changes in a client's health status from one time point to another), P = Prevalence-based measure (measures client's health status at a single point in time)

^c interRAI 1st = interRAI's 1st generation QI set by Hirdes et al. [37], Swiss RAI-HC = Swiss RAI-HC QI set by Burla et al. [29], interRAI 2nd = interRAI's 2nd generation QI set by Morris et al. [30], ✓ = part of the QI set, (✓) = part of the QI set but not core set (only for Swiss QIs)

^d Face validity was assessed for all three QI sets:

A = Face validity assessed with nominal group technique and rated as appropriate

B = Face validity assessed with unstructured consensus process (approved by 70% of experts) and rated as appropriate

C = Rating process of face validity not described but rated as appropriate

D = Face validity assessed with nominal group technique and rated as inappropriate

^e Interrater reliability was tested by Burla et al. [29] for 18 Swiss RAI-HC QIs, values indicate kappa/yules values, N/A = not applicable: interrater reliability not tested due to small sample size. Kappa/yules values < 0.40 were interpreted as insufficient reliability

care. In the absence of QI “gold standard” measurements, construct validity of QIs is often examined instead of criterion validity, e.g. by assessing correlations between process and outcome QIs [36, 48, 49]. However, such assessments were not carried out in the reviewed articles since there are only few process RAI-HC QIs.

The review revealed that the reliability of the identified RAI-HC QIs has hardly been tested so far. For example, interrater reliability of RAI-HC QIs, respectively the underlying RAI-HC items, was only assessed by Burla et al. [29] for a subset of Swiss RAI-HC QIs. Burla et al. [29] found moderate or good interrater reliability for 12 QIs and insufficient interrater reliability for six QIs. Possible reasons for poor interrater reliability are difficulty to detect certain health conditions (e.g. delirium), a high rate of true clinical change and the fluctuations of symptoms and function, misinterpretation of assessment instructions, or poorly designed assessment items [50]. Further, Burla et al. [29] point out that the interrater reliability findings of the 18 QIs must be interpreted cautiously due to the small sample size (only 24 home care clients were assessed). In contrast to the RAI-HC QIs, the psychometric properties of the instrument RAI-HC, i.e. its items and scales, has been tested more thoroughly. These validation studies show overall acceptable reliability results [8, 10, 11, 13]. Nevertheless, good reliability of RAI-HC data does not guarantee

the reliability of QIs based on these data nor the ability to be applied in quality comparisons over time or between providers [51, 52].

In addition, the reported validation process of RAI-HC QIs in the included articles did not involve precision tests to determine the reliability of QIs for distinguishing real differences in performance. In the reviewed studies, between-provider variation was assessed by comparing QI rates (e.g. interquartile range) among HCOs and geographic regions, whereby for interRAI's 1st and 2nd generation QIs unadjusted as well as risk-adjusted rates and for the Swiss RAI-HC QIs only unadjusted rates were computed. However, the possibility that a substantial amount of between-provider variation is attributable to random variation has not been considered. Chance can cause substantial differences in performance in the absence of true quality differences [6]. The empirical evaluation of QIs for the acute care [6, 53–55] and nursing home setting [46, 56–59] includes other statistical methods to test for reliability such as the intra-class correlation coefficient or more advanced modeling techniques, such as multilevel or Bayesian-based estimation as well as thorough risk adjustments. To determine whether RAI-HC QIs have the ability to consistently measure quality differences in home care, the application of the above mentioned statistical procedures should be considered.

Table 3 Methodological characteristics of RAI-HC quality indicator sets (AIRE instrument)

	interRAI's 1st generation QI set [37]	Swiss RAI-HC QI set [29]	interRAI's 2nd generation QI set [30]
Domain 1: Purpose, relevance and organizational context	60%	60%	47%
The purpose of the indicator is described clearly	4	4	3.5
The criteria for selecting the topic of the indicator are described in detail	2.5	4	3
The organizational context of the indicator is described in detail	4	3.5	2
The quality domain the indicator addresses is described in detail	2.5	1.5	2.5
The health-care process covered by the indicator is described and defined in detail	1	1	1
Domain 2: Stakeholder Involvement	44%	28%	56%
The group developing the indicator includes individuals from relevant professional groups	3.5	2.5	4
Considering the purpose of the indicator, all relevant stakeholders have been involved at some stage of the development process	2.5	2	3
The indicator has been formally endorsed	1	1	1
Domain 3: Scientific evidence	11%	0%	0%
Systematic methods were used to search for scientific evidence	1.5	1	1
The indicator is based on recommendations from an evidence-based guideline	1.5	1	1
The supporting evidence has been critically appraised	1	1	1
Domain 4: Additional evidence, formulation and usage	69%	48%	54%
The numerator and denominator are described in detail	4	4	4
The target patient population of the indicator is defined clearly	4	2	2
A strategy for risk adjustment has been considered and described	4	1	4
The indicator measures what it is intended to measure (validity)	2.5	2.5	2.5
The indicator measures accurately and consistently (reliability)	1	3.5	1
The indicator has sufficient discriminative power	3.5	3	3.5
The indicator has been piloted in practice	1	1	1
The efforts needed for data collection have been considered	4	4	4
Specific instructions for presenting and interpreting the indicator results are provided	3.5	1	1.5

Abbreviations: AIRE Appraisal of Indicators through Research and Evaluation, QI Quality indicator

Item scores: Each item score ranges from 1 to 4 with 1 = strongly disagree (confident that the criterion has not been fulfilled or no information was available), 2 and 3 = disagree/agree (unsure whether the criterion has been fulfilled) and 4 = strongly agree (confident that the criterion has been fulfilled) [33]

Domain score calculation: Domain scores for the four AIRE instrument domains were calculated as follows: first, the two authors' scores per item were summed up and divided by two to obtain an average rating per item; second, the average item scores were summed up per domain to obtain the domain score; and third, the domain score were standardized using the following formula: (total score per domain - minimum possible score) / (maximum possible score - minimum possible score) × 100%

High methodological quality of QI set: If score ≥ 50% across all four AIRE instrument domains

Strengths and limitations

To the authors' knowledge, this is the first attempt to identify and summarize, in a systematic way, the scientific evidence on validity and reliability of RAI-HC QIs, thereby identifying gaps for potential improvement in future validation studies. The review was limited by the small number of articles available. While it cannot be ruled out that validation studies regarding RAI-HC QIs may not have been published in peer-reviewed journals, grey literature searches did not provide additional publications. To the best of our knowledge, we have reviewed all published work on the validity and reliability of RAI-HC QIs. Due to the poor reporting of methodology and

results, it is difficult to draw a firm conclusion on the overall validity and reliability of the QIs. Furthermore, the QI assessment with the AIRE instrument was hindered by the limited information in the validation processes of the RAI-HC QI sets. Thus, the AIRE instrument rating results have to be interpreted cautiously.

Conclusion

Based on the description of the RAI-HC QI sets, the validation processes, and the methodological assessment with the AIRE instrument, this review indicates that the quality of the evidence underpinning the identified RAI-

HC QIs is weak and information about validity and reliability is scarce. QIs that are not valid and reliable result in an inaccurate or unreliable measure of the quality of care and, therefore, are neither useful for identifying poor nor desirable quality of care [7]. In addition, information on the methodological quality of QIs is crucial for different stakeholders such as health care providers or policy-makers when selecting QIs for their intended use. This review provides suggestions as to what additional testing of QIs are needed to strengthen their scientific soundness. Considering that RAI-HC QIs are already implemented and frequently used by HCOs for quality improvement processes but also in scientific research to measure and compare home care quality among HCOs or regions, more evidence of the validity and reliability of RAI-HC QIs is essential.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12913-020-05238-x>.

Additional file 1. Search strategy used in PubMed.

Abbreviations

ADL: Activities of daily living; AIP: Agency Intake Profile; AIRE: Appraisal of Indicators through Research and Evaluation; ED: Emergency department; HCO: Home care organization; IADL: Instrumental activities of daily living; MD: Medical doctor; NGT: Nominal group technique; QI: Quality indicator; RAI-HC: Resident Assessment Instrument-Home Care; U.S.: United States; WRHA: the Winnipeg Regional Health Authority

Acknowledgements

We wish to thank Cornelis Kooijman and Esther Baettig from the Swiss Association of Home Care Organizations (Sitex Schweiz) for their support. The study relates to ongoing work by the HCD (HomeCareData) Research Group, who provided input on the results. The HCD Research Group consists of: Institute of Health Sciences, ZHAW: Julia Dratva, René Schaffert, Aylin Wagner. Winterthur Institute of Health Economics, ZHAW: Eva Hollenstein, Florian Liberatore, Sarah Schmelzer. Swiss Health Observatory (OBSAN): Laure Dutoit, Sonia Pellegrini. Institute of Social and Preventive Medicine, University of Bern: Adrian Spoerri, Andreas Boss.

Ethics approval and consent to participation

Not applicable.

Authors' contributions

AW, RS, FZ and JD conceptualized the study design. AW, RS, FZ and NM conducted the systematic literature review. AW and FZ conducted the literature search, AW the data extraction and synthesis of the evidence. AW and NM evaluated the risk of bias of included studies and quality indicator sets. JD, RS, FZ and NM provided analytic support and contributed to the interpretation of findings. AW was responsible for writing the first and all consecutive drafts and the submitted paper. All authors provided critical commentary in the draft revisions according to their expertise. All authors approved the submitted version of the manuscript and agreed to be personally accountable for the accuracy or integrity of any part of the work.

Funding

The study was funded by the Swiss National Science Foundation (SNSF), National Research Program 74 "Smarter Health Care", Project "Swiss Home Care Data: patient profiles and quality measures for home care" (No. 167499);

and the Swiss Federal Office of Public Health (FOPH). The funding bodies had no role in the design nor the execution of the study. Further, they did not participate in analysis or interpretation of the data, nor manuscript writing.

Availability of data and materials

Data sharing is not applicable to this article as no datasets were generated or analysed during the current study.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹ZHAW Zurich University of Applied Sciences, School of Health Professions, Institute of Health Sciences, Technikumstrasse 71, 8401 Winterthur, Switzerland. ²University of Basel, Medical Faculty, Department of Public Health, Institute of Nursing Science, Bernoullistrasse 28, 4056 Basel, Switzerland. ³University of Basel, Medical Faculty, Klingelbergstrasse 61, 4056 Basel, Switzerland.

Received: 18 November 2019 Accepted: 19 April 2020

Published online: 29 April 2020

References

- Colombo F, Llana-Nozal A, Mercier J, Tjadens F. Help wanted? Providing and paying for long-term care. OECD Publishing; 2011. doi:10.1787/9789264097759-en.
- United Nations, Department of Economics and Social Affairs, population division. World population ageing 2017. 2017. <https://www.un.org/en/development/desa/population/theme/ageing/WPA2017.asp>. Accessed 4 Sep 2019.
- Lafortune G, Balestat G. Trends in severe disability among elderly people: assessing the evidence in 12 OECD countries and the future implications. OECD health working papers. Paris: OECD; 2007.
- Lehnert T, Heuchert M, Hussain K, König HH. Stated preferences for long-term care: a literature review. *Ageing Soc* 2019;39:1873–1913. doi:10.1017/S0144686X18000314.
- Mainz J. Defining and classifying clinical indicators for quality improvement. *Int J Qual Health Care* 2003;15:523–530. doi:10.1093/intqhc/mzg081.
- Lovaglio PG. Benchmarking strategies for measuring the quality of healthcare: problems and prospects. *Sci World J* 2012;2012:606154. doi:10.1100/2012/606154.
- Campbell S, Braspenning J, Hutchinson A, Marshall M. Research methods used in developing and applying quality indicators in primary care. *Qual Saf Health Care* 2002;11:358–364. doi:10.1136/qhc.11.4.358.
- Morris JN, Fries BE, Steel K, Ikegami N, Bernabei R, Carpenter GI, et al. Comprehensive clinical assessment in community setting: applicability of the MDS-HC. *J Am Geriatr Soc* 1997;45:1017–1024. doi:10.1111/j.1532-5415.1997.tb02975.x.
- Landi F, Tua E, Onder G, Carrara B, Sgadari A, Rinaldi C, et al. Minimum data set for home care: a valid instrument to assess frail older people living in the community. *Med Care* 2000;38:1184–1190. doi:10.1097/00005650-200012000-00005.
- Hirdes JP, Ljunggren G, Morris JN, Frijters DH, Finne Soveri H, Gray L, et al. Reliability of the interRAI suite of assessment instruments: a 12-country study of an integrated health information system. *BMC Health Serv Res* 2008;8:277. doi:10.1186/1472-6963-8-277.
- Poss JW, Jutan NM, Hirdes JP, Fries BE, Morris JN, Teare GF, et al. A review of evidence on the reliability and validity of minimum data set data. *Healthc Manage Forum* 2008;21:33–39. doi:10.1016/S0840-4704(10)60127-5.
- Kim H, Jung YI, Sung M, Lee JY, Yoon JY, Yoon JL. Reliability of the interRAI long term care facilities (LTCF) and interRAI home care (HC). *Geriatr Gerontol Int* 2015;15:220–228. doi:10.1111/ggi.12330.
- Hogeveen SE, Chen J, Hirdes JP. Evaluation of data quality of interRAI assessments in home and community care. *BMC Med Inform Decis Mak* 2017;17:150. doi:10.1186/s12911-017-0547-9.

14. Morris JN, Fries BE, Mehr DR, Hawes C, Phillips C, Mor V, et al. MDS cognitive performance scale. *J Gerontol* 1994;49:M174–M182. doi:10.1093/geronj/49.4.M174.
15. Hartmaier SL, Sloane PD, Guess HA, Koch GG, Mitchell CM, Phillips CD. Validation of the minimum data set cognitive performance scale: agreement with the mini-mental state examination. *J Gerontol Ser A*. 1995; 50A:M128–M133. doi:10.1093/gerona/50A.2.M128.
16. Morris JN, Fries BE, Morris SA. Scaling ADLs within the MDS. *J Gerontol Ser A*. 1999;54:M546–M553. doi:10.1093/gerona/54.11.M546.
17. Burrows A. Development of a minimum data set-based depression rating scale for use in nursing homes. *Age Ageing* 2000;29:165–172. doi:10.1093/ageing/29.2.165.
18. Fries BE, Simon SE, Morris JN, Flodstrom C, Bookstein FL. Pain in U.S. nursing homes: validating a pain scale for the minimum data set. *The Gerontologist* 2001;41:173–179. doi:10.1093/geront/41.2.173.
19. Leung DY, Leung AY, Chi I. An evaluation of the factor structure of the instrumental activities of daily living involvement and capacity scales of the minimum data set for home care for elderly chinese community dwellers in Hong Kong. *Home Health Care Serv Q* 2011;30:147–159. doi:10.1080/01621424.2011.592421.
20. Penny K, Barron A, Higgins AM, Gee S, Croucher M, Cheung G. Convergent validity, concurrent validity, and diagnostic accuracy of the interRAI depression rating scale. *J Geriatr Psychiatry Neurol* 2016;29:361–368. doi:10.1177/0891988716666376.
21. Salahudeen MS, Nishtala PS. A systematic review evaluating the use of the interRAI home care instrument in research for older people. *Clin Gerontol* 2019;42:463–484. doi:10.1080/07317115.2018.1447525.
22. Quality indicators. interRAI. <http://www.interrai.org/indicators.html>. Accessed 1 Sep 2019.
23. Bos JT, Frijters DHM, Wagner C, Carpenter GI, Finne-Soveri H, Topinkova E, et al. Variations in quality of home care between sites across Europe, as measured by home care quality indicators. *Aging Clin Exp Res* 2007;19:323–329. doi:10.1007/BF03324709.
24. Mofina AM, Guthrie DM. A comparison of home care quality indicator rates in two Canadian provinces. *BMC Health Serv Res* 2014;14:37. doi:10.1186/1472-6963-14-37.
25. Lo AT, Gruneir A, Bronskill SE, Bierman AS. Sex differences in home care performance: a population-based study. *Womens Health Issues* 2015;25: 232–238. doi:10.1016/j.whi.2015.01.004.
26. Foebel AD, van Hout HP, van der Roest HG, Topinkova E, Garms-Homolova V, Frijters D, et al. Quality of care in European home care programs using the second generation interRAI home care quality indicators (HCQIs). *BMC Geriatr* 2015;15. doi:10.1186/s12877-015-0146-5.
27. Joling KJ, van Eenoo L, Vetrano DL, Smaardijk VR, Declercq A, Onder G, et al. Quality indicators for community care for older people: a systematic review. *PLoS One* 2018;13:e0190298. doi:10.1371/journal.pone.0190298.
28. Higgins JPT, Green S. *Cochrane handbook for systematic reviews of interventions*. Chichester, England; Hoboken, NJ: Wiley; 2008.
29. Burla L, Schaffert R, Mylaeus M, Rüesch P. [Developing and testing quality indicators of home care in Switzerland]. *Gesundheitswesen*. 2010;72:106–113. doi:10.1055/s-0029-1233474.
30. Morris JN, Fries BE, Frijters D, Hirdes JP, Steel RK. interRAI home care quality indicators. *BMC Geriatr* 2013;13:127. doi:10.1186/1471-2318-13-127.
31. Harman LE, Guthrie DM, Cohen J, Declercq A, Fisher K, Goodridge D, et al. Potential quality indicators for seriously ill home care clients: a cross-sectional analysis using resident assessment instrument for home care (RAI-HC) data for Ontario. *BMC Palliat Care* 2019;18:3. doi:10.1186/s12904-018-0389-y.
32. Seow H, Harman L, Barbera L, Sutradhar R, McGrail K, Burge F, et al. Pan-Canadian quality indicators for patients at end of life derived from interRAI data. *J Pain Symptom Manag* 2018;56:e59–e60. doi:10.1016/j.jpainsymman.2018.10.174.
33. de Koning J, Smulders A, Klazinga N. The appraisal of indicators through research and evaluation (AIRE) instrument. Amsterdam: Academic Medical Center; 2006.
34. Burkett E, Martin-Khan MG, Gray LC. Quality indicators in the care of older persons in the emergency department: a systematic review of the literature. *Australas J Ageing* 2017;36:286–298. doi:10.1111/ajag.12451.
35. Petrosyan Y, Sahakyan Y, Barnsley JM, Kuluski K, Liu B, Wodchis WP. Quality indicators for care of depression in primary care settings: a systematic review. *Syst Rev* 2017;6:126. doi:10.1186/s13643-017-0530-7.
36. Voeten SC, Krijnen P, Voeten DM, Hegeman JH, Wouters MWJM, Schipper IB. Quality indicators for hip fracture care, a systematic review. *Osteoporos Int* 2018;29:1963–1985. doi:10.1007/s00198-018-4558-x.
37. Hirdes JP, Fries BE, Morris JN, Ikegami N, Zimmerman D, Dalby DM, et al. Home care quality indicators (HCQIs) based on the MDS-HC. *The Gerontologist* 2004;44:665–679. doi:10.1093/geront/44.5.665.
38. Dalby DM, Hirdes JP, Fries BE. Risk adjustment methods for home care quality indicators (HCQIs) based on the minimum data set for home care. *BMC Health Serv Res* 2005;5:7. doi:10.1186/1472-6963-5-7.
39. Kidder D, Rennison M, Goldberg H, Warner D, Bell B, Hadden L, et al. MegaQI covariate analysis and recommendations: identification and evaluation of existing quality indicators that are appropriate for use in long-term care settings. Cambridge, Massachusetts: Abt Associates Inc; 2002.
40. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychol Assess* 1994;6:284–290. doi:10.1037/1040-3590.6.4.284.
41. Center for Health Policy/Center for Primary Care and Outcomes Research & Battelle Memorial Institute. Quality indicator measure development, implementation, maintenance, and retirement. Rockville, MD: Agency for Healthcare Research and Quality; 2011. https://www.qualityindicators.ahrq.gov/Downloads/Resources/Publications/2011/QI_Measure_Development_Implementation_Maintenance_Retirement_Full_5-3-11.pdf.
42. Boulkedid R, Abdoul H, Loustau M, Sibony O, Alberti C. Using and reporting the Delphi method for selecting healthcare quality indicators: a systematic review. *PLoS One* 2011;6:e20476. doi:10.1371/journal.pone.0020476.
43. Linstone HA, Turoff M. *The Delphi method: techniques and applications*. London: Addison-Wesley; 1975.
44. Fitch K, Bernstein SJ, Aguilar MD, Burnand B, LaCalle JR, Lazaro P, et al. The RAND/UCLA appropriateness method user’s manual. Santa Monica: RAND Corporation; 2001.
45. Delbecq AL, van de Ven AH, Gustafson DH. *Group techniques for program planning: a guide to nominal group and delphi processes*. Glenview: Scott, Foresman & Co.; 1975.
46. Jones RN, Hirdes JP, Poss JW, Kelly M, Berg K, Fries BE, et al. Adjustment of nursing home quality indicators. *BMC Health Serv Res* 2010;10:96. doi:10.1186/1472-6963-10-96.
47. Nakrem S, Vinsnes AG, Harkless GE, Paulsen B, Seim A. Nursing sensitive quality indicators for nursing home care: international review of literature, policy and practice. *Int J Nurs Stud* 2009;46:848–857. doi:10.1016/j.ijnurstu.2008.11.005.
48. Fischer C, Lingsma HF, Anema HA, Kievit J, Steyerberg EW, Klazinga N. Testing the construct validity of hospital care quality indicators: a case study on hip replacement. *BMC Health Serv Res* 2016;16. doi:10.1186/s12913-016-1778-7.
49. Gooiker GA, Kolfschoten NE, Bastiaannet E, Velde CJH van de, Eddes EH, Harst E van der, et al. Evaluating the validity of quality indicators for colorectal cancer care. *J Surg Oncol* 2013;108:465–471. doi:10.1002/jso.23420.
50. Hirdes JP, Smith TF, Rabinowitz T, Yamauchi K, Pérez E, Telegdi NC, et al. The resident assessment instrument-mental health (RAI-MH): inter-rater reliability and convergent validity. *J Behav Health Serv Res* 2002;29:419–432. doi:10.1007/BF02287348.
51. Berg K, Mor V, Morris J, Murphy KM, Moore T, Harris Y. Identification and evaluation of existing nursing homes quality indicators. *Health Care Financ Rev*. 2002;23:19–36.
52. Hutchinson AM, Milke DL, Maisey S, Johnson C, Squires JE, Teare G, et al. The resident assessment instrument-minimum data set 2.0 quality indicators: a systematic review. *BMC Health Serv Res* 2010;10:166. doi:10.1186/1472-6963-10-166.
53. Fung V, Schmittiel JA, Fireman B, Meer A, Thomas S, Smider N, et al. Meaningful variation in performance: a systematic literature review. *Med Care* 2010;48:140–148. doi:10.1097/MLR.0b013e3181bd4dc3.
54. Selby JV, Schmittiel JA, Lee J, Fung V, Thomas S, Smider N, et al. Meaningful variation in performance: what does variation in quality tell us about improving quality? *Med Care* 2010;48:133–139. doi:10.1097/MLR.0b013e3181c15a6e.
55. van Dishoeck AM, Lingsma HF, Mackenbach JP, Steyerberg EW. Random variation and rankability of hospitals using outcome indicators. *BMJ Qual Saf* 2011;20:869–874. doi:10.1136/bmjqs.2010.048058.
56. Berlowitz DR, Christiansen CL, Brandeis GH, Ash AS, Kader B, Morris JN, et al. Profiling nursing homes using bayesian hierarchical modeling. *J Am Geriatr Soc* 2002;50:1126–1130. doi:10.1046/j.1532-5415.2002.50272.x.

57. Castle NG, Degenholtz H, Engberg J. State variability in indicators of quality of care in nursing facilities. *J Gerontol Ser A* 2005;60:1173–1179. doi:10.1093/gerona/60.9.1173.
58. Arling G, Kane RL, Lewis T, Mueller C. Future development of nursing home quality indicators. *The Gerontologist* 2005;45:147–156. doi:10.1093/geront/45.2.147.
59. Arling G, Lewis T, Kane RL, Mueller C, Flood S. Improving quality assessment through multilevel modeling: the case of nursing home compare. *Health Serv Res* 2007;42:1177–1199. doi:10.1111/j.1475-6773.2006.00647.x.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

