

SYSTEMATIC REVIEW

Open Access

Measuring team factors thought to influence the success of quality improvement in primary care: a systematic review of instruments

Sue E Brennan^{1*}, Marije Bosch², Heather Buchan³ and Sally E Green¹

Abstract

Background: Measuring team factors in evaluations of Continuous Quality Improvement (CQI) may provide important information for enhancing CQI processes and outcomes; however, the large number of potentially relevant factors and associated measurement instruments makes inclusion of such measures challenging. This review aims to provide guidance on the selection of instruments for measuring team-level factors by systematically collating, categorizing, and reviewing quantitative self-report instruments.

Methods: Data sources: We searched MEDLINE, PsycINFO, and Health and Psychosocial Instruments; reference lists of systematic reviews; and citations and references of the main report of instruments. Study selection: To determine the scope of the review, we developed and used a conceptual framework designed to capture factors relevant to evaluating CQI in primary care (the InQulRe framework). We included papers reporting development or use of an instrument measuring factors relevant to teamwork. Data extracted included instrument purpose; theoretical basis, constructs measured and definitions; development methods and assessment of measurement properties. Analysis and synthesis: We used qualitative analysis of instrument content and our initial framework to develop a taxonomy for summarizing and comparing instruments. Instrument content was categorized using the taxonomy, illustrating coverage of the InQulRe framework. Methods of development and evidence of measurement properties were reviewed for instruments with potential for use in primary care.

Results: We identified 192 potentially relevant instruments, 170 of which were analyzed to develop the taxonomy. Eighty-one instruments measured constructs relevant to CQI teams in primary care, with content covering teamwork context (45 instruments measured enabling conditions or attitudes to teamwork), team process (57 instruments measured teamwork behaviors), and team outcomes (59 instruments measured perceptions of the team or its effectiveness). Forty instruments were included for full review, many with a strong theoretical basis. Evidence supporting measurement properties was limited.

Conclusions: Existing instruments cover many of the factors hypothesized to contribute to QI success. With further testing, use of these instruments measuring team factors in evaluations could aid our understanding of the influence of teamwork on CQI outcomes. Greater consistency in the factors measured and choice of measurement instruments is required to enable synthesis of findings for informing policy and practice.

Keywords: Continuous quality improvement, Primary care, Evaluation, Systematic review, Measurement, Instrument, Conceptual framework, Taxonomy, Teamwork, Team functioning

* Correspondence: sue.brennan@monash.edu

¹School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia

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

Background

The use of cross-functional teams to diagnose process-based quality problems, and develop and test process improvements, is an important element of continuous quality improvement (CQI) [1-3]. Cross-functional teamwork aims to capitalize on the varied knowledge and perspectives of team members, encouraging collaboration that is expected to lead to better problem solving, more innovative decisions, and greater engagement in implementing proposed solutions [4-6]. Cross-functional teams may also promote organizational learning [7]. In CQI, the use of cross-functional teams is expected to lead to better processes of care and greater adherence to the new processes [2,3,8]. Realizing these benefits requires teams with both task-specific competencies (knowledge and skills required to use CQI methods) and teamwork competencies (knowledge and skills that enable members to function as an effective team). It also requires a context that enables teams to overcome well-documented barriers to cross-functional teamwork, such as professional boundaries and status differences that hinder collaboration [4,6,9-11].

Despite the centrality of teamwork to CQI, the extent to which team functioning influences CQI outcomes is not well understood [12-14]. In a systematic review of contextual factors thought to influence QI success, thirteen studies examined team-level factors [14]. Team leadership, team climate, team process, and physician involvement in the QI team appeared to be important, but the supporting evidence was scant [14]. Research on QI collaboratives—an approach in which teams use CQI methods to introduce change—has led to a greater focus on the influence of teams on QI outcomes (e.g., [15-20]). Yet in a recent review of the impact of QI and safety teams [13], the authors concluded that existing studies provided limited information about the attributes of successful QI teams or factors that influence team success.

One of the challenges to addressing the limited evidence base is the variability in how team-level factors are conceptualized and measured in QI studies [14]. This variability fragments the evidence base, making it difficult to compare and synthesize findings across studies. Recent calls to address these challenges focus on the need for theory development to explain how CQI works and factors that influence its effectiveness, and the identification of valid and reliable measures to enable theories to be tested [21-26].

In this paper, we report a systematic review of instruments measuring team-level factors thought to influence the success of CQI. This review is part of a larger project aiming to aid the evaluation of CQI in primary care by providing guidance on factors to include in evaluations and the selection of instruments for measuring these factors. The project includes a companion review of instruments measuring organizational, process, and individual-level factors [27], and development of a

conceptual framework, the Informing Quality Improvement Research (InQuIRe) in primary care framework.

Our initial framework is included in this paper (Figure 1) to illustrate the scope of the review of instruments and as the basis for assessing the coverage of available instruments. The framework reflects our initial synthesis of CQI theory, a summary of which is reported in the companion review of instruments [27]. In brief, we identified recurrent themes about the core components of CQI from landmark papers that stimulated adoption of CQI in healthcare (e.g., [2,28-31]). We built on these themes, refining and adding concepts from the main bodies of CQI research and models of CQI (e.g., [32-35]). We searched for models, frameworks, and theories intended to describe how context influences quality improvement and change (e.g., [36-39]). From these sources, we extracted factors salient to primary care, grouping them thematically using existing models to guide categorization. Our analysis was augmented by frameworks for understanding implementation, learning, and innovation (e.g., [40-43]) and models of teamwork (e.g., [12,44-47]). Both sources contributed new factors and informed the final categorization and structure of our framework. Our analysis of instruments is used to integrate new factors and concepts into the framework. These refinements are reported in the taxonomies presented in the reviews of instruments. The final InQuIRe framework will be reported separately.

This review aims to provide guidance on the selection of instruments for measuring team-level factors in studies of CQI in primary care. The specific objectives are to: identify self-report instruments measuring team-level factors thought to modify the effect of CQI; determine how the factors measured have been conceptualized; develop a taxonomy for categorizing instruments based on our initial framework and new concepts arising from the review of instruments; use the taxonomy to categorize and compare the content of instruments, enabling assessment of the coverage of instruments suitable for evaluating CQI team function in primary care; appraise the methods of development and testing of existing instruments, and summarize evidence of their validity, reliability, and feasibility for measurement in primary care. We focus on self-report instruments because of their utility in quantitative studies examining the relationship between team context, team process, and outcomes. Alternative measurement methods, such as behavioral observation scales, require substantial resources and are generally not feasible for larger scale evaluation.

Our review differs from two existing reviews of instruments measuring teamwork [48,49] in its focus on QI in primary care. The existing reviews consider all forms of teamwork in healthcare, including a large number of instruments relevant only to clinical teams [48,49]. QI teams perform tasks and interact in ways that differ markedly from clinical work [50,51]. For example, collaborative

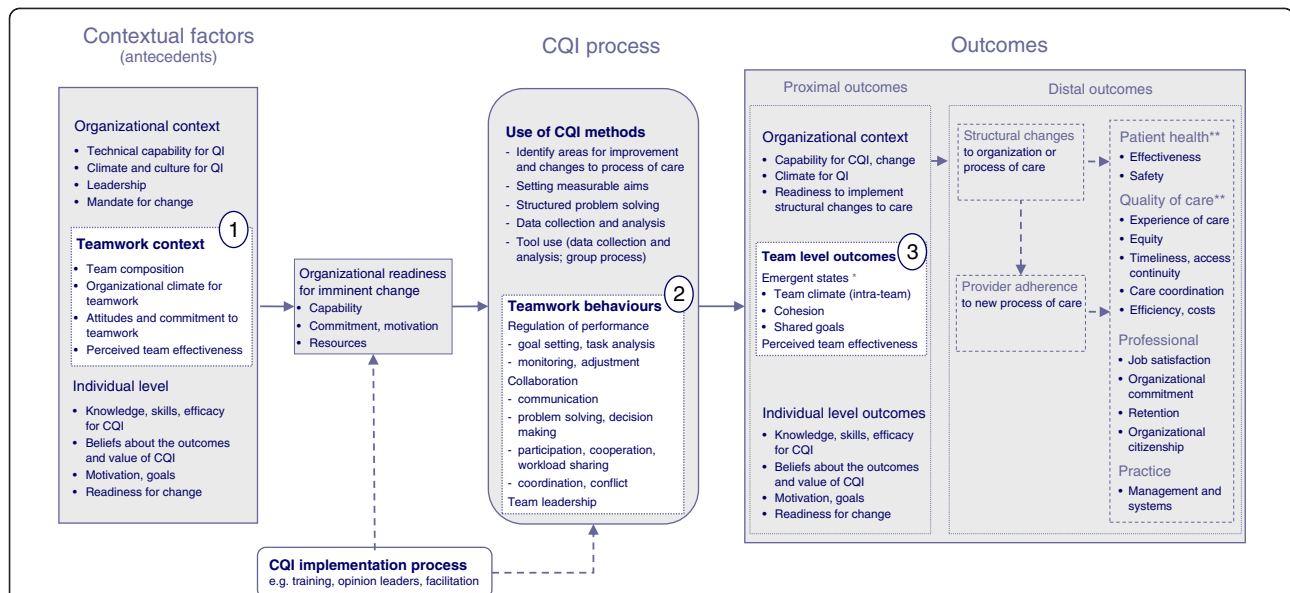


Figure 1 Conceptual framework for defining the scope of the review – initial version of Informing Quality Improvement Research (InQuIRe) in primary care. Instruments within the scope of the review reported in this paper cover three domains (shaded in white and numbered as follows in the figure and throughout the review): (1) Teamwork context; (2) Team process; (3) Proximal team outcomes. Areas shaded in grey are covered in a companion review. Those in boxes with dashed lines are outside the scope of either review. * Emergent states are perceptions and capabilities of the team that arise from interaction between team members. Although depicted as a proximal outcomes, in subsequent cycles of teamwork they become antecedents that contribute to the teamwork context. ** These outcomes are based on dimensions of quality from Institute of Medicine (U.S.) Committee on Quality of Health Care in America. *Crossing the quality chasm: a new health system for the 21st century.* Washington DC: National Academy Press, 2001:xx, 337.

problem solving in QI involves eliciting views on the underlying cause of problems, questioning existing ways of working, and integrating different perspectives to identify changes to care. In primary care, there may be particular challenges and risks in engaging these behaviors, especially in small practices where the same team works closely to deliver and improve care. Team climate (e.g., trust, open communication, norms of decision-making) may have a heightened role in influencing team function [47], and practices often lack the broader organizational structures and resources that facilitate QI work [52-54]. By identifying instruments suitable for primary care, we aim to help researchers measure factors salient to understanding QI teamwork in this understudied context [55].

Scope of the review: InQuIRe framework

This review covers instruments relevant to three domains of the InQuIRe framework: teamwork context, team process, and proximal team outcomes (shaded in white and numbered one to three in Figure 1). Teamwork context encompasses organizational, team, and individual factors thought to influence how the CQI team functions [44,46,51,56]. Within this domain, organizational climate for teamwork reflects shared perceptions of the extent to which the practice supports and rewards CQI teamwork through its policies, practices, procedures, and behavioral

expectations [57]. Team process captures interactions between team members during the CQI process, focusing on behaviors required for members to function as an effective CQI team. These behaviors include knowledge sharing, collaborative problem solving, and making full use of members’ perspectives [45,58]. Proximal outcomes (emergent states and perceived team effectiveness) result from interactions between team members [44-46]. These outcomes include development of a shared understanding of team goals, commitment to achieving these goals, and perceptions of whether the climate within the team is safe for engaging in the behaviors required of CQI teams. In subsequent cycles of teamwork, these factors become antecedents that contribute to the teamwork context. Figure 2 illustrates terms used to describe the framework with an example from the taxonomy (see Additional file 1 for a glossary).

Methods

Figure 3 summarizes the four stages of the review and lists the criteria for selection of instruments at each stage.

Stage one: searching and initial screening

Data sources and search methods

We searched MEDLINE (from 1950 through October 2011), PsycINFO (from 1967 through October 2011),

and Health and Psychosocial Instruments (HaPI) (from 1985 through February 2012) using controlled vocabulary (thesaurus terms and subject headings) and free-text terms. Reference lists of identified systematic reviews were screened. Snowballing techniques (citation searches, reference lists) were used to trace the development and use of instruments included in Stage four. Searches were limited to articles published in English. Search terms and details of the search strategy are reported in Additional file 2. Abstracts and the full text of potentially relevant studies were screened for inclusion by one author (SB). Papers reporting development or use of a quantitative, self-report instrument measuring factors within the scope of the team domains of the InQuIRE framework were included for data extraction (see Figure 3 for initial inclusion criteria).

Stage two: development of taxonomy for categorizing instrument content

Data extraction

One review author extracted data from all studies for all stages of the review (SB). To refine the data extraction guidance and data extraction, a research assistant extracted data from a sub-sample of studies (15 papers,

comprising a 10% sample of the studies included in all stages of review). Across the two companion reviews double data extraction was performed on 30 papers. Data extracted for Stage two are described in Table 1. Data included descriptions of the instrument purpose and format, and data to facilitate analysis and categorisation of the content of each instrument (e.g., constructs measured; theoretical basis).

Taxonomy development

Methods for developing the taxonomy were based on the framework approach for qualitative data analysis [59]. This approach combines deductive methods (commencing with concepts and themes from an initial framework) with inductive methods (based on themes that emerge from the data). The InQuIRE framework (Figure 1) provided the initial structure and content for our taxonomy. We used content analysis of instruments to refine the taxonomy, aiming to ensure comprehensive coverage of relevant factors.

Instruments confirmed as relevant to one or more of the three domains of our framework were included for content analysis. To capture the breadth of potentially relevant constructs, we included instruments irrespective of whether item content was suitable for primary care (i.e., items that did not resonate in primary care settings or inferred the team worked within a large organization, e.g., 'I would accept almost any job in order to keep working with this team' [60]). The content of each instrument (items, sub-scales), and associated construct definitions, was compared with the taxonomy. Instrument content that matched constructs in the taxonomy was summarized using existing labels. The taxonomy was expanded to include missing constructs and new concepts, initially using the labels and descriptions reported by the instrument developers.

To ensure the taxonomy was consistent with the broader literature, we reviewed definitions extracted from review articles and conceptual papers identified from the search. We also searched for and used additional sources to define constructs when included studies did not provide a definition, a limited number of studies contributed to the definition, or the definition provided appeared inconsistent with our initial concept or that in other included studies. Following analysis of all instruments and supplementary sources, related constructs were grouped in the taxonomy. Overlapping constructs were collapsed, distinct constructs were assigned a label reflecting the broader literature, and the dimensions of constructs were specified to create the taxonomy.

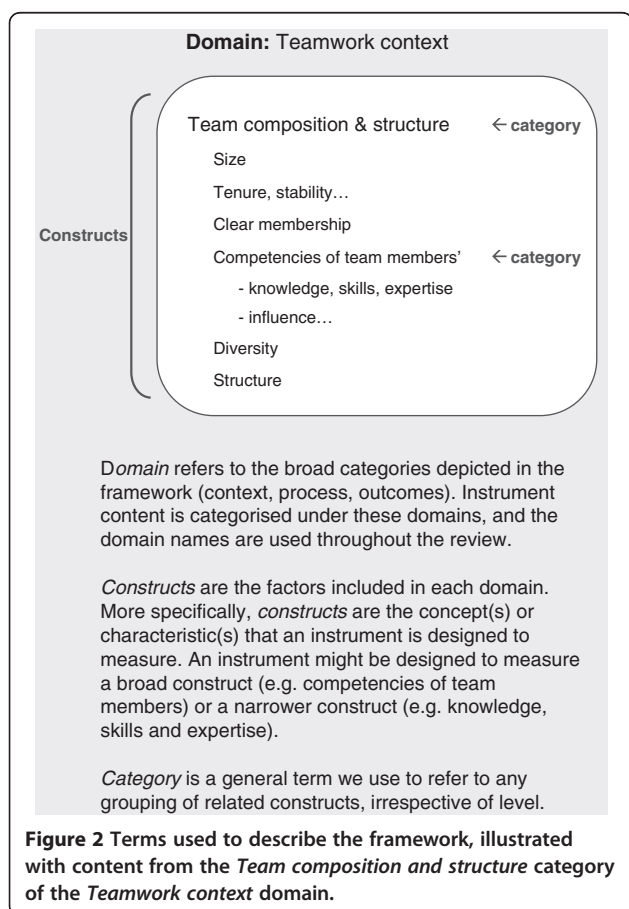
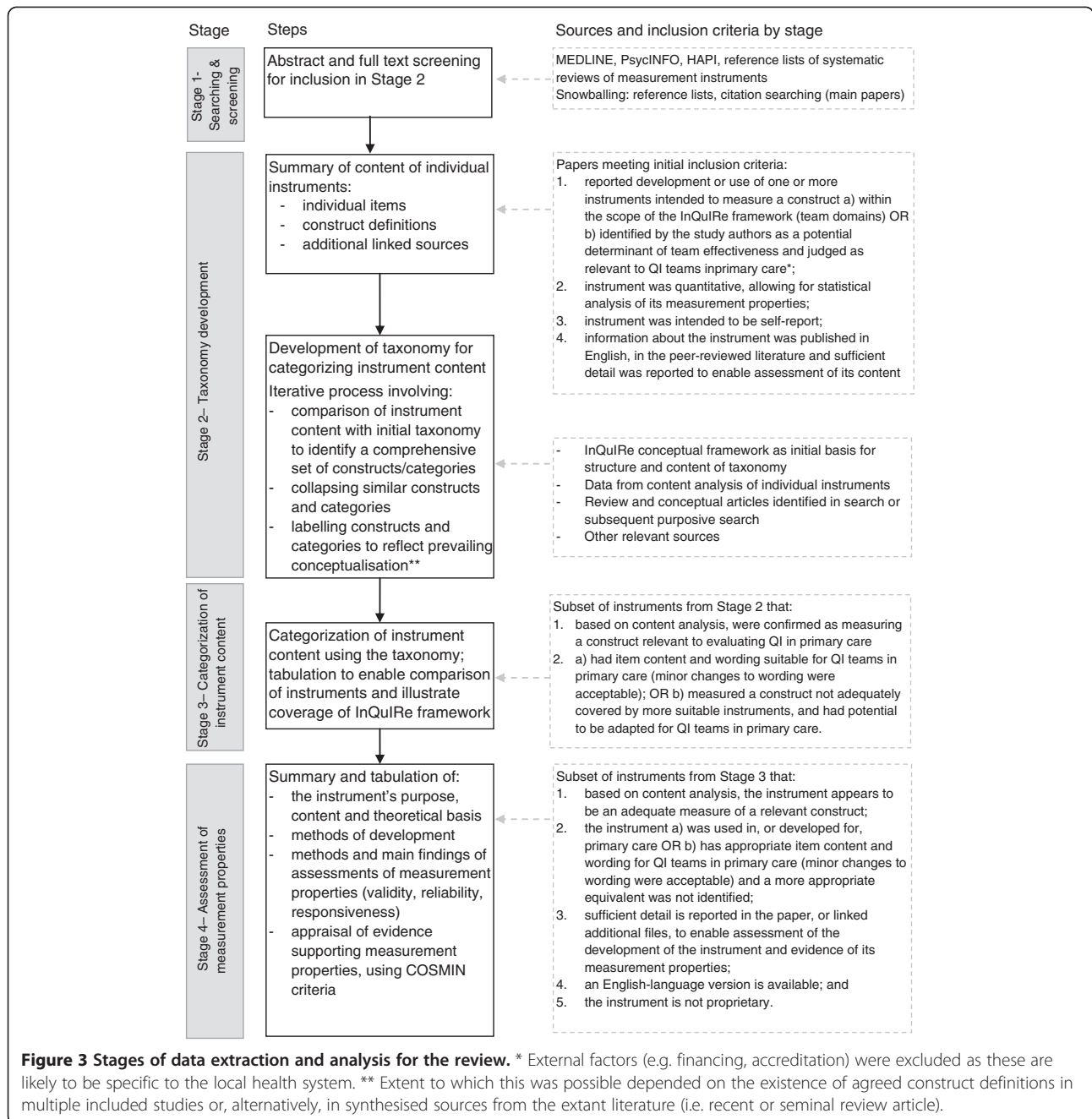


Figure 2 Terms used to describe the framework, illustrated with content from the *Team composition and structure* category of the *Teamwork context* domain.

Stage three: categorization of instrument content

One author (SB) categorized the content of all instruments confirmed as measuring a relevant construct with



item content and wording suitable for QI teams in primary care (Figure 3, sources and inclusion criteria for Stage 3). For constructs not adequately covered by suitable instruments, we included instruments with potential for adaptation (e.g., instruments with a strong theoretical basis designed to measure attitudes toward clinical teamwork). Categorization of instrument content was based on the final set of items reported in the main report(s) for each instrument. A second author (MB) independently categorized the content of a sub-sample of ten instruments (12%), including all instruments where

there was any uncertainty over categorization. The categorization was discussed to identify revisions to the taxonomy and confirm final categorization. All instruments were then re-categorized using the final taxonomy.

Stage four: assessment of measurement properties

Information about the development and assessment of measurement properties of each instrument was extracted from the main and secondary reports (the latter focusing on studies of QI or change in primary care) (Table 2). These data included descriptions of the methods

Table 1 Data extracted at stage two

Data extracted	Description
Study characteristics	Study aims Study design (categorized as experimental, observational, instrument development, model development) Setting in which the instrument was used
Instrument source	Name of instrument Source paper for the instrument as cited by the authors
Instrument purpose	Purpose for which the instrument was used (descriptive, predictive or diagnostic, outcome measure/evaluative)
Instrument format	Number of items Response scale (Likert, ipsative, etc.); response options
Instrument content and theoretical basis	Constructs and dimensions measured Definitions of the constructs; additional description of the content required to illustrate how the construct had been operationalised (<i>e.g.</i> , sample items) Theoretical basis of the instrument and references cited for the theory

and findings of assessments of content and construct validity, reliability, and acceptability of the instrument to respondents.

We used the COSMIN (COnsensus-based Standards for the selection of health status Measurement Instruments) checklist [61] to appraise the methods used for development and testing of instruments. The COSMIN criteria are intended for studies reporting measures of patient reported outcomes, however the checklist has

strong evidence for its content validity [61,62] and mirrors the Joint Committee on Standards for Educational and Psychological Testing [63] (indicating relevance to measures other than health outcomes). Minor changes were incorporated based on guidance from organizational psychology [64-66], and reviews of organizational measures (*e.g.*, [67-69]).

Most instruments had undergone limited testing, and reporting of information required to complete the checklist

Table 2 Data extracted at stage four

Data extracted	Description ¹
Instrument development	Methods used to generate items (<i>e.g.</i> , items derived from existing instruments; new items generated from data from interviews or comprehensive review of theory) Methods used to refine instrument
Administration & scoring	Method of administration (<i>e.g.</i> , self-administered, facilitated) Feasibility of administration (<i>e.g.</i> , researcher time, resources) Acceptability to respondents (<i>e.g.</i> , views on burden and complexity) Methods of scoring and analysis
Measurement properties	Methods and findings of assessments of: Content validity (<i>e.g.</i> , clear description of content domain and theoretical basis, expert assessment of items for relevance and comprehensiveness) Construct validity - Hypothesis testing (<i>e.g.</i> , whether scores on the instrument converge with measures of theoretically related variables, discriminate between groups, predict relevant outcomes) - Instrument structure (<i>e.g.</i> , using factor analytic methods) Reliability (<i>e.g.</i> , internal consistency, stability over time, inter-rater) Responsiveness
Other assessments	Interpretability (potential for ceiling and floor effects; guidance on what constitutes an important change or difference in scale scores) Generalizability (sampling methods, description of sample, and response rate reported)

¹Definitions of the extracted measurement properties are provided in Additional file 3.

was sparse. Because of the sparse data, for each instrument we tabulated a summary of the extent of evidence available for each property, and a description of the instrument's development and testing. We used appraisal data to provide an overall summary of the methods used to develop and test each instrument.

Results

Screening and identification of unique instruments

Our search identified 5,629 non-duplicate references, 5,178 of which were excluded following abstract screening (Figure 4). We reviewed the full text of 451 articles, and included 306 papers in stage two (Figure 4). Of these, 151 reported studies in healthcare, 49 of which were in primary care. Fifty-five papers had as their primary aim development of an instrument. Observational studies were most commonly reported in papers ($n = 214$), encompassing descriptive studies and tests of theoretical models. Twenty-one papers reported experimental designs. The remaining papers were conceptual. Individual papers reported between one and four potentially relevant instruments, contributing 339 reports of development or use of 170 unique instruments (Figure 4).

Development of taxonomy and categorization of instrument content

The content of 170 instruments was analyzed to develop the taxonomy. The analysis led to elaboration and extension of the constructs in our initial framework, as reflected in the revised taxonomy (Additional file 4: Tables S3-S5). The main modifications are described under 'Content and coverage of the domains of the InQuIRE framework.' Minor changes were incorporated based on discussion arising from independent categorization of the subset of instruments. These changes involved adding lower level categories to some factors. There was consensus on most categories and no changes to the structure of the taxonomy.

We categorized the content of 81 instruments, of which 45 measured aspects of organizational context, 57 measured team process, and 59 measured proximal team outcomes (some instruments covered more than one domain, hence the total sums to >81). We excluded 89 instruments because items were unsuitable for QI teams in primary care or the authors reported example items only (see Figure 4 for examples [47,70-72]). The categorization of instrument content is reported by domain of the InQuIRE framework (Additional file 4: Tables S3-S5). This serves as a guide to content, however instruments vary substantially in comprehensiveness and the strength of their theoretical basis. In the next section, we review the instruments most suitable for measuring each domain. Used in combination with the results tables (Additional files 5: Table S7 and 6: Table S8), this analysis can be used

to select instruments and to identify areas where better instruments are required.

Content and coverage of domains of the InQuIRE framework

1) Teamwork context - enabling conditions and attitudes to teamwork

Table S3 (Additional file 4) reports the final taxonomy and categorization of instrument content for the teamwork context domain (box number '1' in the InQuIRE framework, Figure 1).

Description of the teamwork context domain We included instruments in this domain if they measured perceptions of task design, team composition and structure, or organizational climate for teamwork. These three categories comprise enabling conditions, a term adopted from Hackman's model of team effectiveness [73,74]. The conceptualization of enabling conditions reflected in the taxonomy builds on models of team effectiveness (especially, [1,12,75]). Task design is an addition to the initial taxonomy, included to capture views about the nature of the team task (*e.g.*, perceptions of the need for individuals to work interdependently [75]).

We included individual attitudes, beliefs, and values as part of teamwork context. These factors reflect individual views about teamwork in general. They are distinct from emergent states, which are perceptions and cognitions of the team that emerge from interactions between members [44,45]. To reflect this, emergent states (and perceived team effectiveness) are categorized as proximal outcomes (box '3' in the InQuIRE framework, Figure 1). In subsequent cycles of teamwork, these proximal outcomes contribute to teamwork context.

Comprehensive measures of enabling conditions

Wageman 2005 [73] and Campion 1993 [76,77] are the most comprehensive instruments measuring enabling conditions, covering aspects of task design (*e.g.*, interdependence, autonomy), team composition (*e.g.*, competencies, diversity), and organizational climate (*e.g.*, recognition, training). Neither instrument was developed in healthcare, nor have they had much prior use in healthcare contexts. Shortell 2004 [78] was the most comprehensive instrument developed for healthcare QI teams. It includes items measuring task design and organizational climate.

Measures of task design We identified multiple instruments measuring aspects of task design (*e.g.*, Van den Bossche 2006a [79] and Van der Vegt 1998 [80] measure interdependence; Kirkman 1999a [81] and Langfred 2005 [82] measure autonomy; Edmondson 1999c [83] measures task significance). Items measuring perceptions

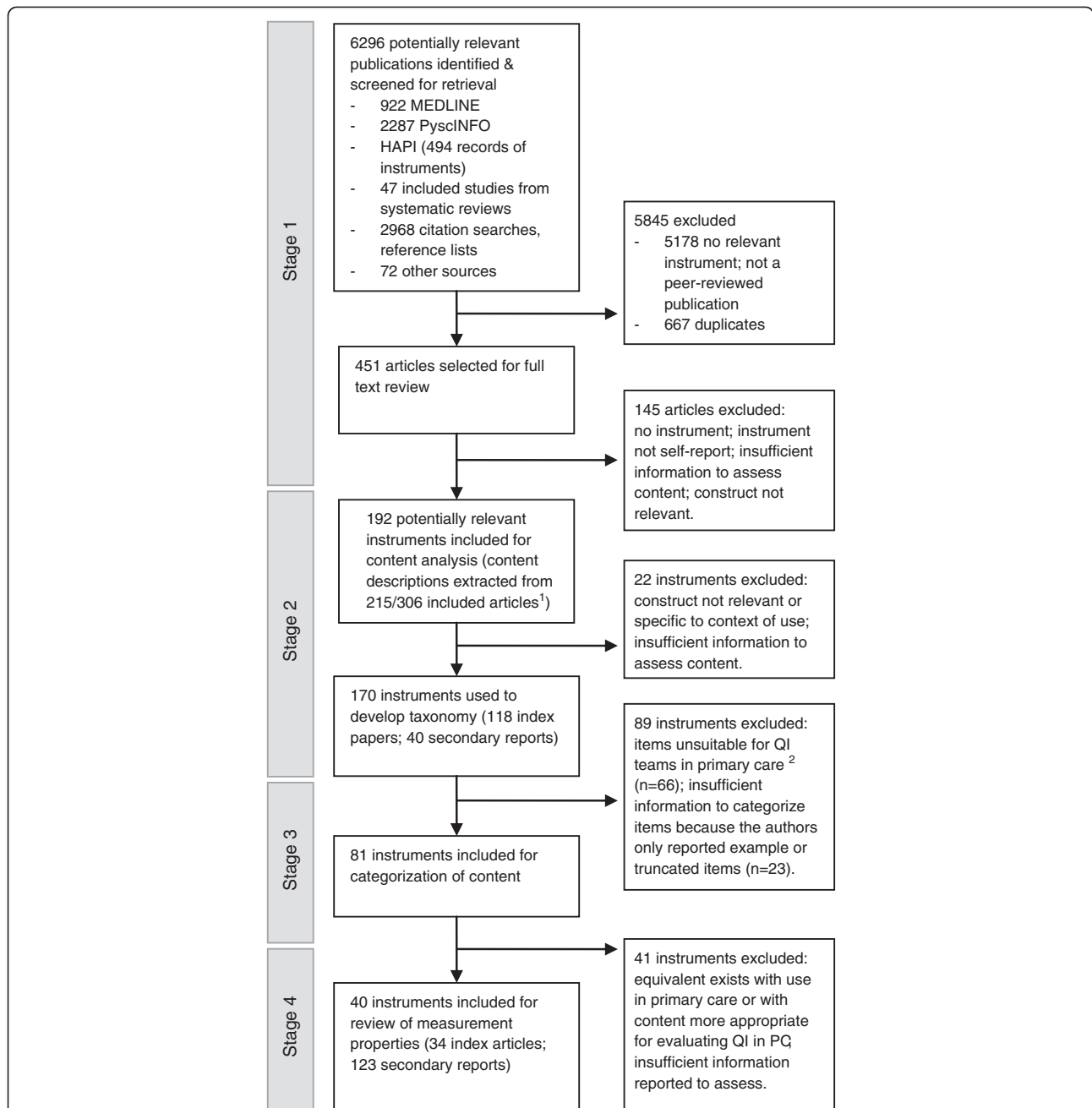


Figure 4 Flow of studies and instruments through the review. ¹Remainder of 306 articles (n=91) were secondary reports that did not contribute additional information about instrument content. These were retained for assessment of measurement properties if required when final set of studies for inclusion in Stage 4 was determined. ²Instruments considered unsuitable were those (i) with content intended for a specific context of use (e.g., Poulton's measure of team effectiveness has multiple items specific to the UK National Health Service [47]; Schroder's collaborative practice assessment tool is intended for clinical care teams [70]), (ii) with content adequately covered by more suitable instruments (e.g., measures of transformational leadership (e.g. [71]) were excluded because we identified multiple measures of leadership in relation to teamwork), and (iii) instruments that, on further analysis, were not self-report measures (e.g., Irvine's team problem solving effectiveness scale requires document analysis [72]).

of task meaningfulness or significance were relatively common in instruments developed for healthcare QI teams (e.g., Mills 2004 [19], Alemi 2001 [84], Lukas 2009 [85], Schouten 2010 [86]).

Measures of team composition and structure Few instruments measured perceptions of team composition. Those identified focused on perceptions of members' task-related skill and knowledge (e.g., Mills 2004 [19], Edmondson 1999c [83], Millward 2001 [87]). Diversity was typically measured using indices created from other team composition variables (e.g., knowledge, status) rather than perception-based measures [11]. We identified two potentially relevant instruments measuring team structure. Bunderson 2010a [88] measures specialization (distinct and specialist roles), hierarchy (clear leadership), and formalization (goals and scheduling). Thylefors 2005 [89] measures team structure related to cross-functional clinical team effectiveness. It requires adaptation for QI teams.

Measures of organizational climate for teamwork Shortell 2004 [78] and Lukas 2009 [85] include items measuring organizational climate for QI teamwork in healthcare, with an emphasis on external leadership support. The salience of leader support for QI was evident in instruments designed for healthcare QI teams; most included items measuring this construct (e.g., Lemieux-Charles 2002a [8], Lubomski 2008 [90], Mills 2004 [19], Schouten 2010 [86]). Instruments measuring external leadership were included if they measured support for teamwork. Of these, instruments measuring support for team self-management were common, however items require adaptation because they reflect the manufacturing origin of this construct (e.g., Manz 1987 [91]).

Measures of individual attitudes, beliefs and values about teamwork These instruments typically examine views about team process and outcomes (e.g., Dobson 2009a [92]), or beliefs about individual capability for teamwork (e.g., Fulmer 2005 [93], Eby 1997 [94]). Most instruments from healthcare measured views about clinical teamwork. We included those with greatest potential for adaptation (e.g., Fulmer 2005 [93], Baker 2010 [95], Heinemann 1999 [96]). Adaptation would involve rewording and formal assessment of the relevance of scales and items to QI teams. Scales measuring preferences for individual versus collective work (individualism/collectivism) indicate whether individuals value teamwork (e.g., Driskell 2010 [97], Shaw 2000 [98]).

2) Team process (teamwork behaviors)

Tables S4a and S4b (Additional file 4) report the final taxonomy and categorization of instrument content

within the team process domain (box number '2' in the InQuIRE framework, Figure 1).

Description of the team process domain Team processes are the 'interactions that take place among team members' to organize 'task-work to achieve collective goals' [45] and combine 'cognitive, motivational/affective, and behavioral resources' [44]. We included instruments in this domain if they measured teamwork behaviors or interactions between team members. We adapted existing frameworks for team processes [45] and teamwork behaviors [45,58] to develop categories within this domain. In the resulting taxonomy, instrument content was categorized as measuring regulation of team performance, collaborative behaviors and interpersonal processes, team maintenance, learning behaviors, and team leadership behaviors.

Regulation of team performance involves behaviors closely reflected in the CQI process. These include goal specification and planning, monitoring task performance, and adjusting to ensure goal attainment [58]. Monitoring teamwork process, feedback, and reflection are other important regulatory behaviors. Collaborative behaviors encompass the social processes that enable teams to develop shared knowledge and understanding, capitalize on diverse knowledge and perspectives, and work collaboratively to achieve goals [44,45,58,79,99]. Team maintenance was added to our taxonomy to delineate behaviors that enhance team viability (e.g., conflict management, motivational behaviors). Most instruments measuring team process covered aspects of regulation, collaboration, and maintenance, hence they are described under a general heading of 'team process.'

Measures of learning behaviors and team leadership behaviors were distinct from other instruments measuring process (so are described separately). Learning behaviors were an addition to our taxonomy, defined as 'an ongoing process of reflection and action, characterized by asking questions, seeking feedback, experimenting, reflecting on results, and discussing errors or unexpected outcomes of actions.' ([83], p353). Team leadership includes behaviors of designated and informal leaders, the latter encompassing shared leadership and when individuals '... emerge informally as a leader' [100].

Measures of team process developed for QI teams in healthcare Schouten 2010 [86] focuses on regulation of performance through goal specification, planning, and monitoring progress toward goals. Alemni 2001 [84] covers similar concepts, but the response format limits psychometric analysis. Shortell 2004 [78] also emphasizes CQI methods used to plan and test changes, focusing on collaborative decision-making. Items were based on Lemieux-Charles 2002a [8], which measures

collaborative behaviors during CQI. Wilkens 2006 [101] includes stand-alone scales measuring feedback and conflict. Other measures are tailored to specific projects or provide narrower coverage of constructs (see Lukas 2009 [85], Mills 2004 [19], Irvine 2002 [72], Duckers 2008 [18], Lubomski 2008 [90]).

Broad measures of team process Hoegl 2001 [102] is a comprehensive measure of collaborative behaviors, with stand-alone scales measuring communication, coordination, workload sharing, and effort. Thompson 2009 [103] provides good coverage of collaborative problem solving and decision making behaviors. Hiller 2006 [104] measures the extent to which teams share responsibility for problem solving, goal specification, and planning. Anderson 1998 [105] focuses on team climate (an emergent state), but includes items measuring regulatory and collaborative behaviors thought to support innovation.

Scales measuring single constructs and complex aspects of team process Stand-alone scales measuring specific team process constructs (e.g., individual scales measuring decision making) were uncommon in healthcare. Instruments from non-healthcare settings address this gap. Relevant scales include measures of group process during meetings (Kuhn 2000 [106]), the extent to which teams reflect on their performance (reflexivity) (Schippers 2007 [107], Brav 2009 [108], de Jong 2010b [109]), information exchange and knowledge sharing (Bunderson 2010b [88], Staples 2008a [110]), cooperation (Brav 2009 [108]), and effort (de Jong 2010b [109]). Several instruments from non-healthcare settings address complex aspects of interpersonal processes, drawing on substantive bodies of theoretical and empirical research. Exemplars include Janssen 1999 [111] and Tjosvold 1986 [112]. These instruments measure constructive controversy behaviors that enable teams to draw out and integrate divergent knowledge and perspectives, through 'skilled discussion' of opposing views ([112], p127).

Measures of learning behaviors Three instruments explicitly measured learning behaviors (Savelsbergh 2009 [113], van den Bossche 2006b [79], Edmondson 1999a [83]). Savelsbergh 2009 is a comprehensive measure of reflection, feedback, and communication behaviors that enable team members to develop a shared understanding of problems. Both Savelsbergh 2009 and van den Bossche 2006b build on theories of the role of discourse in developing shared cognitions [79,114]. Some instruments include direct questions about learning (e.g., 'Members of my team actively learn from one another' Mathieu 2006 [115]), but are not measures of behaviors that support learning.

Measures of team leadership behaviors Despite the perceived importance of leadership in teams, few instruments incorporated measures of leadership behavior. Lemieux Charles 2002a [8] is the most comprehensive measure for QI teams. Other relevant instruments cover shared leadership (e.g., Hiller 2006 [104]), leader behaviors that empower teams (e.g., Arnold 2000 [116]), and coaching (e.g., Wageman 2005 [73]).

3) Proximal team outcomes emergent states and perceived team effectiveness

Tables S5a and S5b (Additional file 4) reports the final taxonomy and categorization of instrument content within the proximal team outcomes domain (box number '3' in the InQuIRE framework, Figure 1).

Description of the proximal team outcomes domain

Instruments with content measuring emergent states or perceived team effectiveness were included in this domain. We adopted Marks' view of emergent states as dynamic properties of a team that result from interaction between members. These proximal outcomes of team process include team-level cognitions, motivations, values, and beliefs [44-46]. Our taxonomy specifies common emergent states, focusing on those most pertinent to QI teams. These are categorized as beliefs about team capability, team knowledge, empowerment, cohesion, trust, team identification and commitment, team climate, and team norms.

We adopt three categories of perceived team effectiveness from Cohen and Bailey [75], and Hackman [73,74]: task performance outcomes, attitudinal outcomes (e.g., satisfaction with the team, team viability), and behavioral outcomes (i.e., changes to teamwork capability).

Multi-dimensional measures of emergent states including team climate

Climate is a multi-dimensional construct that typically focuses on a specific behavior (e.g., climate for learning). Scales measuring individual emergent states can be combined to measure the dimensions of climate (or team context) most relevant to the target behavior. Anderson 1998 [105,117,118] is the most widely used example. Focusing on team climate for innovation, Anderson 1998 includes scales measuring innovation, quality orientation, psychological safety, and shared vision. Other examples include instruments measuring team context for QI (Wilkens 2006 [101]), context for team learning (e.g., Edmondson 1999b [83]; van den Bossche 2006a [79]), and climate for creativity (Barczak 2010 [119]; Caldwell 2003 [120]). Each of these instruments contains scales suitable for measuring single emergent states. Millward 2001 [87] is a more general measure of factors thought to influence team effectiveness (knowledge, identification, cohesion).

Scales measuring single emergent state constructs

Many scales exist for measuring single emergent states, most from non-healthcare settings. We focused on those measuring factors inadequately covered by multi-dimensional instruments. Examples include instruments measuring empowerment (Kirkman 1999a [81]), cohesion (Carless 2000a [121]), team trust (Costa 2011 [122]), team identification and commitment (Janssen 2008 [123]; Bishop [124]), learning orientation (Bunderson 2003 [125]), and transactive memory systems (a team's shared awareness of and ability to access its members' knowledge) (Lewis 2003 [126]).

Measures of emergent states developed for QI teams in healthcare

Emergent states were measured in instruments developed for healthcare QI teams; however, items were often combined in a scale with items measuring process (e.g., Shortell 2004 [78], Schouten 2010 [86], Mills 2004 [19]; Lemieux Charles 2002a [8]; Lubomski 2008 [90]). In general, this makes them unsuitable for measuring specific emergent states.

Measures of perceived team effectiveness Most instruments measuring perceived effectiveness were written for the study in which they were used, with limited information about the basis for their content. We focused on dimensions of effectiveness that could not be measured objectively (e.g., willingness to continue with the team) rather than perceptions of process of care or clinical outcomes. Of the instruments used in healthcare QI teams, Lemieux-Charles 2002b [8] includes items measuring satisfaction and viability, and Irvine 2000b [127] measures perceived success.

Instrument characteristics, development, and measurement properties

In Stage four, we reviewed the development and measurement properties of 40 instruments with use or potential for use in primary care (unshaded in Tables S3 to S5, Additional file 4). This information is intended to aid selection of instruments based on evidence of reliability, validity, feasibility of administration, and acceptability to respondents.

We summarize characteristics of each included instrument (e.g., purpose, dimensions as described by developers, items, and response scale) and provide examples of use relevant to CQI evaluation (Additional file 5). The characteristics of excluded instruments (shaded in grey in Tables S3 to S5, Additional file 4) and reasons for exclusion are reported in Additional file 6. Instruments are grouped by setting (as per Additional file 4: Tables S3 to S5), and then reported in alphabetical order.

Table S6 (Additional file 4) provides an overview of the development and testing of measurement properties

for each instrument. It indicates the extent of evidence in the main report(s) and studies in relevant contexts, and is based on assessments described in Additional file 7: Table S9.

Most reports of instrument development included comprehensive definitions that reflected related research and theory. Reports of instruments from the psychology literature (25 of 40 instruments) provided this detail more frequently than reports from healthcare, and included more comprehensive descriptions of the intended measurement domain. Reports of instruments from healthcare varied widely in how comprehensively they described the intended measurement domain, and some of these instruments appeared to have no theoretical basis. Reporting of the process used to develop items was scant. Only a quarter of studies reported an independent assessment of content validity (e.g., using an expert consensus process).

For most instruments, evidence of construct validity (e.g., through hypothesis testing, analysis of the instrument's structure) was derived from one or two studies (exceptions include Shortell 1991 [128], Jehn 2008a [99], Campion 1993 [76]). Few instruments had evidence that they were predictive of objective measures of team effectiveness, and only six instruments had evidence that they can differentiate between groups. Across all the criteria, Anderson 1998 (Team Climate for Innovation inventory) was an exception [105,117,118]. With extensive use and testing, it is the only included instrument with multiple tests of construct validity in healthcare (some in primary care). These include tests of whether the measure predicts relevant outcomes and differentiates between groups.

Most studies report Cronbach's alpha (a measure of internal consistency or the 'relatedness' between items) for the scale (or subscales); however, this was sometimes done without checks to ensure that the scale was unidimensional [129,130]. About one-half of the studies assessed some form of inter-rater reliability, to determine within-team agreement. This property should be assessed when inferences are made at team level from individual-level data.

Most reports covered conceptual and analytical issues associated with measuring collective constructs, at a minimum specifying the level at which the construct was defined and interpreted. In most cases, analytical methods were appropriate for the level at which inferences were to be made.

Few studies reported the potential for floor and ceiling effects (which may influence ability to detect change in a construct [130]), and none of the studies provided guidance on what constitutes an important change in scores. The acceptability of the instrument to respondents was reported for less than a quarter of instruments. Missing

items, assessment of whether items were missing at random or due to other factors, and the potential for response bias [130,131] were rarely reported.

Discussion

In this review, we aim to provide guidance for researchers seeking to measure team-level factors that potentially influence the process and outcomes of CQI. We identified many potentially useful instruments and scales, with some novel examples from the psychology literature. Collectively, these instruments cover many of the factors hypothesized to contribute to QI success. Inclusion of these measures in CQI evaluations could help address some key questions about the extent to which team function and the context in which teams work influence CQI outcomes. Many of the included instruments had little or no prior use in healthcare settings, especially in primary care, and some instruments had limited evidence supporting their measurement properties. Additional testing of the measurement properties of most instruments in relevant contexts is therefore required. We consider these issues, and the potential application of available instruments, in relation to each domain of the InQuIRe framework. We then discuss key considerations for researchers in relation to the use and development of instruments.

Measurement of teamwork context

Enabling conditions provide teams with the structure, resources, and broader organizational support required to do their work [73,74]. Attributes of the task itself provide the motivation for team members to work together and invest effort in achieving team goals. These conditions are important for CQI teams, and there are a broad range of instruments available for their measurement. Among these instruments are many that include short scales (two to three items) that could be used to measure single constructs. Combining these scales to measure factors salient to the context in which evaluations are conducted is the most efficient way to measure aspects of team context. Some of the constructs we included in our taxonomy are almost entirely absent from the QI literature, yet their inclusion could add to our understanding of CQI team function. For example, perceptions of interdependence may be an important determinant of active participation on a CQI team. While individuals may be supportive of teamwork in general (*i.e.*, display positive attitudes toward teamwork), they may believe that their contribution on a CQI team will go unrecognized (outcome interdependence), will not be influential in achieving the team goals (task interdependence), or will not help them achieve personal goals (goal interdependence) [79,80].

Measurement of team process

Measures of team process are underutilized in studies of CQI, yet they have the potential to illuminate the extent to which teams enact formal teamwork behaviors and the extent to which these behaviors account for variance in CQI outcomes. A team's ability to capitalize on its collective knowledge and expertise rests on its ability to use effective collaborative and interpersonal processes. Measures of team process are a measure of the fidelity with which the 'teamwork' component of CQI is implemented. They provide important explanatory data about why CQI interventions may work in some contexts and not others. They can also provide data for developing targeted interventions to enhance a team's ability for CQI teamwork. Although there are a number of instruments that are potentially useful, there is an absence of evidence about the extent to which these self-report instruments measure actual behavior. Criterion measures for behavior involve observing teams and rating their behavior on a validated scale (*e.g.*, van Ginkel's scale for rating 'elaboration' during team discussion [132]). Behavioral observation scales were excluded from the review because they require substantial resources (trained observers, some requiring analysis of qualitative data). An important area of inquiry would be to establish the construct validity of self-report measures in studies using observation scales as a criterion measure.

Measurement of proximal team outcomes (emergent states, perceived team effectiveness)

The measurement of team climate for innovation - an emergent state - is widespread in healthcare, largely due to the popularity of Anderson and West's Team Climate for Innovation (TCI) inventory [105]. It has a strong theoretical basis, and the body of evidence supporting the measurement properties of this instrument is unrivalled. It is relevant to CQI, yet we found no examples of its use in this context. The dimensions it measures are used to predict a team's ability to innovate. There are many other emergent states of relevance to QI teams, and we encourage researchers to consider the multitude of short scales available for measuring other factors that shape the interpersonal context in which QI teams work. Instruments that measure context for team learning are a good example (*e.g.*, Edmondson 1999b [83], Van den Bossche 2006a [79]).

Measures of perceived team effectiveness (especially those measuring perceived task outcomes relating to process of care and patient outcomes) have a somewhat limited scope for application in healthcare where objective measures of outcomes are desirable. However, measures of satisfaction with team process, team viability, and improved team functioning are important. These outcomes may have implications for the delivery of care in primary care settings where the same people work together on QI and care

teams. A recent study in primary care suggested that positive experiences of QI teamwork contributed to building teamwork capacity for delivering clinical care [133]. Team building is an important outcome of CQI that may enhance the effectiveness of care teams.

Key considerations for researchers

Researchers using this review to select instruments for use in CQI projects or evaluations will need to weigh up the strengths of different instruments, taking into consideration suitability for the hypotheses they intend to test, acceptable respondent burden, and the availability of evidence supporting measurement properties in settings and conditions pertinent to their study. There is no single instrument suited to all purposes. Given the limited evidence of measurement properties for many instruments, the relevance of instrument content and length are key decision points. Surveys of health professionals typically have low response rates [134], which may introduce bias. Shorter instruments have been shown to achieve higher response rates [135]; however, as yet there is insufficient evidence to determine optimal questionnaire length to maximize response rate [134].

The review findings highlight two other areas that need careful consideration: ensuring conceptual clarity when selecting, interpreting, and developing instruments, and increasing consistency in how team-level factors are conceptualized in QI studies.

Ensuring conceptual clarity in measurement

Marks *et al.* describe the common practice of ‘intermingling’ items measuring emergent states with those measuring team processes. In their view, this leads to ‘serious construct contamination’ [45]. Construct contamination can occur when items measuring different constructs are included in a single scale and was apparent in some instruments we reviewed (*e.g.*, combining an item measuring process ‘the team gathers data from patients’ with an item measuring knowledge ‘the team is familiar with measurement’). Construct contamination makes instruments difficult to interpret, and may prevent synthesis of resulting evidence because it is unclear what construct is being measured. Our categorization reflects the content of the items in each instrument (Additional file 4: Tables S3 to S5), not how items are combined to form a scale (the latter is considered under ‘coverage of the framework’). Statistical tests of an instrument’s structure provide empirical evidence about whether items are related. However, a scale can be formed using factor analysis that does not measure a distinct or meaningful construct. Ideally, the scale’s content validity will be supported by a clear construct definition, *a priori* hypotheses about the relationship between items, and independent assessment of content. In the absence of such information, checking

that items in a scale (or subscale) appear to measure a specific construct can guide selection and interpretation of instruments.

Increasing consistency in how team-level factors are conceptualized

While researchers have raised concerns about variability in how factors thought to influence CQI outcomes have been defined and measured [14], our findings suggest that many team-level factors are clearly and consistently defined in the teamwork literature arising from psychology. Moreover, many team concepts are underpinned by a body of empirical research and theory (*e.g.*, [44,46]). Adoption of these concepts in studies of CQI could accelerate our ability to develop an evidence base that is cumulative, enabling findings to be compared and synthesized across studies.

Strengths and limitations

Although other reviews of teamwork instruments exist [48,49], to our knowledge, this is the first review to systematically collate and categorize instruments available for measuring team-level factors in an evaluation of CQI. Our broad, systematic search of the health and psychology literature enabled us to identify a diversity of potentially relevant instruments. Categorizing the content of included instruments using the developed taxonomy enabled direct comparison of instruments. The taxonomy provides a common language for describing instruments and the factors they measure, addressing some of the complexities researchers face when selecting instruments. By basing the taxonomy on a theoretically-based framework for evaluating CQI, we provide guidance on how the identified instruments could be included in an evaluation.

We used a broad search, but may have missed some instruments. We excluded instruments published in books, the grey literature (*e.g.*, theses), and proprietary sources, and those for which items were not reported in full in published sources. Our rationale for these exclusions was that neither the instruments nor information about their measurement properties is readily accessible to researchers. We excluded papers published in languages other than English. One author extracted data from the majority of studies. This approach was based on checks on a subset of studies that confirmed that there were no important discrepancies between two data extractors. One author developed the taxonomy and categorized the content of instruments (SB) with independent categorization from a second author (MB) for a subset of instruments, including those involving more complex decisions. Given the judgment required to make these decisions, alternative categorizations are possible. This is the first application of our taxonomy, and refinement is likely, especially as new

research emerges on the factors that influence QI and change in primary care.

Conclusions

Cross-functional teams are a core element of CQI, yet there is a paucity of evidence about how team-level factors influence the outcomes of CQI. In this review, we aimed to provide guidance to address some of the challenges researchers face in incorporating these factors in evaluations, particularly around the choice of measurement instruments for use in primary care. The conceptualization and measurement of factors that influence CQI is inherently complex, and primary care settings have unique features that heighten the importance of some factors. Yet, there is a need to address these complexities so that we can understand how teamwork influences CQI success and develop and test strategies to optimize team function. Although individual studies can make an important contribution, synthesis of multiple studies in different contexts is needed to identify the factors that contribute to successful CQI. To ensure that individual studies contribute to cumulative knowledge, consistency in the definition and measurement of factors is required. Such consistency enables comparison and synthesis of findings across studies, and underpins our ability to provide guidance for policy and practice on the implementation and outcomes of CQI.

Additional files

Additional file 1: Glossary of terms used in the review.

Additional file 2: Search strategy and search terms.

Additional file 3: Definition of measurement properties.

Additional file 4: Tables reporting content of instruments measuring teamwork context (Table S3), team process (Tables S4a and b), and proximal team outcomes (Tables S5a and b), and overview of instrument development and assessment of measurement properties (Table S6).

Additional file 5: Table summarizing the characteristics of instruments included for review of measurement properties (Stage four) (Table S7).

Additional file 6: Table summarizing the characteristics of instruments included for categorization of content, but excluded from review of measurement properties (Stage four) (Table S8).

Additional file 7: Table summarizing the development and measurement properties of instruments included in Stage four of the review (Table S9).

Competing interests

Heather Buchan is a member of the Implementation Science Editorial Board. The authors have no other competing interests.

Authors' contributions

SB and SG conceived the study with input from HB. SB designed the review, conducted the searching, screening, data extraction and analysis. SG, HB and MB provided input on the design, provided comment on the analysis and the presentation of results. SB drafted the manuscript and made subsequent revisions. All authors provided critical review of the manuscript. All authors read and approved the final manuscript.

Acknowledgements

This project was funded by a Monash University Faculty of Medicine Strategic Grant and SB was supported by a Monash University Departmental doctoral scholarship. We are grateful to Matthew Page for providing research assistance in piloting the data extraction methods and Katherine Beringer for retrieving papers for inclusion in the review. Finally, we gratefully acknowledge the peer reviewers, Sara Singer and Miranda Laurant, for their very helpful comments on the manuscript.

Author details

¹School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia. ²Central Clinical School, Monash University and National Trauma Research Institute, Melbourne, Australia. ³Australian Commission on Safety and Quality in Health Care (ACSQHC), Sydney, Australia.

Received: 13 June 2012 Accepted: 11 February 2013

Published: 14 February 2013

References

1. Hackman J: **Total quality management: empirical, conceptual and practical issues.** *Adm Sci Q* 1995, **40**:309–329.
2. Berwick D, Godfrey BA, Roessner J: **Curing health care: new strategies for quality improvement: a report on the National Demonstration Project on Quality Improvement in Health Care.** 1st edition. San Francisco: Jossey-Bass; 1990.
3. McLaughlin CP, Kaluzny AD: **Continuous quality improvement in health care: theory, implementation and applications.** 3rd edition. Gaithersburg, Md: Aspen Publishers Inc; 2006.
4. Mitchell R, Parker V, Giles M, White N: **Toward realizing the potential of diversity in composition of Interprofessional health care teams: An examination of the cognitive and psychosocial dynamics of Interprofessional collaboration.** *Med Care Res Rev* 2010, **67**:3–26.
5. van Knippenberg D, De Dreu CKW, Homan AC: **Work group diversity and group performance: An integrative model and research agenda.** *J Appl Psychol* 2004, **89**:1008–1022.
6. Fay D, Borrill C, Amir Z, Haward R, West MA: **Getting the most out of multidisciplinary teams: A multi-sample study of team innovation in health care.** *J Occup Organ Psychol* 2006, **79**:553–567.
7. Ilgen DR, Hollenbeck JR, Johnson M, Jundt D: **Teams in organizations: from Input-Process-Output models to IMO models.** *Annu Rev Psychol* 2005, **56**:517–543.
8. Lemieux-Charles L, Murray M, Baker G, Barnsley J, Tasa K, Ibrahim S: **The effects of quality improvement practices on team effectiveness: a mediational model.** *J Organ Behav* 2002, **23**:533–553.
9. Nembhard IM, Edmondson AC: **Making it safe: the effects of leader inclusiveness and professional status on psychological safety and improvement efforts in health care teams.** *J Organ Behav* 2006, **27**:941–966.
10. Lichtenstein R, Alexander JA, McCarthy JF, Wells R: **Status differences in cross-functional teams: effects on individual member participation, job satisfaction, and intent to quit.** *J Health Soc Behav* 2004, **45**:322–335.
11. Harrison DA, Klein KJ: **What's the difference? Diversity constructs as separation, variety, or disparity in organizations.** *Acad Manage Rev* 2007, **32**:1199–1228.
12. Lemieux-Charles L, McGuire WL: **What do we know about health care team effectiveness? A review of the literature.** *Med Care Res Rev* 2006, **63**:263–300.
13. White DE, Straus SE, Stelfox HT, Holroyd-Leduc JM, Bell CM, Jackson K, Norris JM, Flemons WW, Moffatt ME, Forster AJ: **What is the value and impact of quality and safety teams? A scoping review.** *Implement Sci* 2011, **6**:97.
14. Kaplan HC, Brady PW, Dritz MC, Hooper DK, Linam WM, Froehle CM, Margolis P: **The influence of context on quality improvement success in health care: a systematic review of the literature.** *Milbank Q* 2010, **88**:500–559.
15. Schouten LMT, Hulscher MEJL, Akkermans R, Van Everdingen JJE, Grol RPTM, Huijsman R: **Factors that influence the stroke care team's effectiveness in reducing the length of hospital stay.** *Stroke* 2008, **39**:2515–2521.
16. Tucker AL, Nembhard IM, Edmondson AC: **Implementing New Practices: An Empirical Study of Organizational Learning in Hospital Intensive Care Units.** *Manage Sci* 2007, **53**:894–907.

17. Chin MH, Cook S, Drum ML, Jin L, Guillen M, Humikoski CA, Koppert J: **Improving diabetes care in Midwest community health centres with the health disparities collaborative.** *Diabetes Care* 2004, **27**:2–8.
18. Duckers MLA, Wagner C, Groenewegen PP: **Developing and testing an instrument to measure the presence of conditions for successful implementation of quality improvement collaboratives.** *BMC Health Serv Res* 2008, **8**:172.
19. Mills PD, Weeks WB: **Characteristics of successful quality improvement teams: lessons from five collaborative projects in the VHA.** *Jt Comm J Qual Saf* 2004, **30**:152–162.
20. Baker GR, King H, MacDonald JL, Horbar JD: **Using organizational assessment surveys for improvement in neonatal intensive care.** *Pediatrics* 2003, **111**:e419–425.
21. Batalden P, Davidoff F, Marshall M, Bibby J, Pink C: **So what? Now what? Exploring, understanding and using the epistemologies that inform the improvement of healthcare.** *BMJ Qual Saf* 2011, **20**(Suppl 1):i99–i105.
22. Eccles MP, Armstrong D, Baker R, Cleary K, Davies H, Davies S, Glasziou P, Iltott I, Kinmonth AL, Leng G, et al: **An implementation research agenda.** *Implement Sci* 2009, **4**:18.
23. Robert Wood Johnson Foundation: *Advancing the science of continuous quality improvement.* Princeton, NJ: Robert Wood Johnson Foundation; 2008. <http://www.rwjf.org/en/research-publications/research-features/evaluating-CQI.html>.
24. Walshe K: **Understanding what works—and why—in quality improvement: the need for theory-driven evaluation.** *Int J Qual Health Care* 2007, **19**:57–59.
25. Øvretveit J: **Understanding the conditions for improvement: research to discover which context influences affect improvement success.** *BMJ Qual Saf* 2011, **20**(Suppl 1):i18–i23.
26. Baker GR: **Strengthening the contribution of quality improvement research to evidence based health care.** *Qual Saf Health Care* 2006, **15**:150–151.
27. Brennan S, Bosch M, Buchan H, Green S: **Measuring organizational and individual factors thought to influence the success of quality improvement in primary care: a systematic review of instruments.** *Implement Sci* 2012, **7**:121.
28. Berwick D: **Continuous improvement as an ideal in health care.** *N Engl J Med* 1989, **320**:53–56.
29. Kaluzny AD, McLaughlin CP, Kibbe DC: **Continuous quality improvement in the clinical setting: enhancing adoption.** *Qual Manag Health Care* 1992, **1**:37–44.
30. Laffel G, Blumenthal D: **The case for using industrial quality management science in health care organizations.** *JAMA* 1989, **262**:2869–2873.
31. Batalden PB, Stoltz PK: **A framework for the continual improvement of health care: building and applying professional and improvement knowledge to test changes in daily work.** *Jt Comm J Qual Improv* 1993, **19**:424–447. discussion 448–452.
32. Blumenthal D, Kilo CM: **A report card on continuous quality improvement.** *Milbank Q* 1998, **76**:625–648. 511.
33. Shortell SM, Bennett CL, Byck GR: **Assessing the impact of continuous quality improvement on clinical practice: What it will take to accelerate progress.** *Milbank Q* 1998, **76**:593–624.
34. Boaden R: *G Harvey, C Moxham, and N Proudlove: Quality Improvement: theory and practice in healthcare.* Coventry: NHS Institute for Innovation and Improvement/Manchester Business School; 2008.
35. Powell A, Rushmer R, Davies H: *A systematic narrative review of quality improvement models in health care.* Scotland Edinburgh: NHS Quality Improvement; 2009.
36. O'Brien JL, Shortell SM, Hughes EF, Foster RW, Carman JM, Boerstler H, O'Connor EJ: **An integrative model for organization-wide quality improvement: lessons from the field.** *Qual Manag Health Care* 1995, **3**:19–30.
37. Solberg LI: **Improving medical practice: a conceptual framework.** *Ann Fam Med* 2007, **5**:251–256.
38. Lukas CV, Holmes SK, Cohen AB, Restuccia J, Cramer IE, Shwartz M, Charns MP: **Transformational change in health care systems: an organizational model.** *Health Care Manage Rev* 2007, **32**:309–320.
39. Cohen D, McDaniel RR Jr, Crabtree BF, Ruhe MC, Weyer SM, Tallia A, Miller WL, Goodwin MA, Nutting P, Solberg LI, et al: **A practice change model for quality improvement in primary care practice.** *J Healthc Manag* 2004, **49**:155–168. discussion 169–170.
40. Greenhalgh T, Robert G, MacFarlane F, Bate P, Kyriakidou O: **Diffusion of Innovations in Service Organizations: Systematic Review and Recommendations.** *Milbank Q* 2004, **82**:581–629.
41. Gustafson DH, Hundt AS: **Findings of innovation research applied to quality management principles for health care.** *Health Care Manage Rev* 1995, **20**:16–33.
42. Orzano AJ, McInerney CR, Scharf D, Tallia AF, Crabtree BF: **A knowledge management model: Implications for enhancing quality in health care.** *J Am Soc Inf Sci Technol* 2008, **59**:489–505.
43. Rushmer R, Kelly D, Lough M, Wilkinson JE, Davies HT: **Introducing the Learning Practice—I. The characteristics of Learning Organizations in Primary Care.** *J Eval Clin Pract* 2004, **10**:375–386.
44. Kozlowski SWJ, Ilgen DR: **Enhancing the effectiveness of work groups and teams.** *Psychol Sci Public Interest* 2006, **7**:77–124.
45. Marks MA, Mathieu JE, Zaccaro SJ: **A temporally based framework and taxonomy of team processes.** *Acad Manage Rev* 2001, **26**:356–376.
46. Mathieu J, Maynard TM, Rapp T, Gilson L: **Team effectiveness 1997–2007: A review of recent advancements and a glimpse into the future.** *J Manage* 2008, **34**:410–476.
47. Poulton BC, West MA: **The determinants of effectiveness in primary health care teams.** *J Interprof Care* 1999, **13**:7–18.
48. Heinemann GD, Zeiss AM: *Team performance in health care: assessment and development.* New York, NY: Kluwer Academic/Plenum Publishers; 2002.
49. Valentine MA, Nembhard IM, Edmondson AC: *Measuring teamwork in health care settings: A review of survey instruments.* Boston, MA: Harvard Business School; 2012.
50. Fargason CA, Haddock CC: **Cross functional, integrative team decision-making: Essentials for effective QI in health care.** *Qual Rev Bull* 1992, **7**:157–163.
51. Fried B, Carpenter WL: **Understanding and improving team effectiveness in quality improvement.** In *Continuous quality improvement in health care: theory, implementation and applications.* 3rd edition. Edited by McLaughlin C, Kaluzny A. Gaithersburg, Md: Aspen Publishers Inc; 2006:154–190.
52. Solberg LI, Hroszkowski MC, Sperl-Hillen JM, Harper PG, Crabtree BF: **Transforming medical care: case study of an exemplary, small medical group.** *Ann Fam Med* 2006, **4**:109–116.
53. Crabtree BF, Nutting PA, Miller WL, McDaniel RR, Stange KC, Roberto Jaen C, Stewart E: **Primary Care Practice Transformation Is Hard Work: Insights From a 15-Year Developmental Program of Research.** *Med Care* 2011, **49**:10–26.
54. Orzano AJ, Tallia AF, Nutting PA, Scott-Cawiezell J, Crabtree BF: **Are attributes of organizational performance in large health care organizations relevant in primary care practices?** *Health Care Manage Rev* 2006, **31**:2–10.
55. Alexander JA, Hearld LR: **What can we learn from quality improvement research? A critical review of research methods.** *Med Care Res Rev* 2009, **66**:235–271.
56. Fried B, Topping S, Rundall T: **Groups and teams in health services organizations.** In *Health care management: Organization design and behavior.* Edited by Shortell SM, Kaluzny AD. Albany, NY: Delmar; 2006:174–211.
57. Schneider B, Ehrhart MG, Macey WH: **Perspectives on Organizational Climate and Culture.** In *APA handbook of industrial and organizational psychology*, Volume 1. Edited by Zedeck S. Washington, DC: American Psychological Association; 2011.
58. Rousseau V, Aube C, Savoie A: **Teamwork behaviors - A review and an integration of frameworks.** *Small Gr Res* 2006, **37**:540–570.
59. Pope C, Mays N: *Qualitative research in health care.* 3rd edition. Oxford, UK; Malden, Massachusetts: Blackwell Publishing/BMJ Books; 2006.
60. Bishop JW, Dow Scott K: **An examination of organizational and team commitment in a self-directed team environment.** *J Appl Psychol* 2000, **85**:439–450.
61. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, de Vet HCW: **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**:737–745.
62. Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, Bouter LM, de Vet HC: **Protocol of the COSMIN study: COnsensus-based Standards for the selection of health Measurement INstruments.** *BMC Med Res Methodol* 2006, **6**:2.
63. Joint Committee on Standards for Educational and Psychological Testing (U.S.), American Educational Research Association, American Psychological Association, Education. NCoMi: *Standards for educational and psychological testing.* Washington, DC: American Educational Research Association; 1999.
64. Hinkin T: **A brief tutorial on the development of measures for use in survey questionnaires.** *Organ Res Methods* 1998, **1**:104–121.

65. Klein KJ, Kozlowski SWJ: **From micro to meso: Critical steps in conceptualizing and conducting multilevel research.** *Organ Res Methods* 2000, **3**:211–236.
66. Malhotra MK, Grover V: **An assessment of survey research in POM: from constructs to theory.** *J Oper Manag* 1998, **16**:407–425.
67. Holt DT, Armenakis AA, Harris SG, Feild HS: **Toward a comprehensive definition of readiness for change: a review of research and instrumentation.** In *Research in organizational change and development*, Volume 16. Edited by Pasmore WA, Woodman RW. Bingley, UK: Emerald Group Publishing Limited; 2007:289–336.
68. Mannion R, Davies H, Scott T, Jung T, Bower P, Whalley D, McNally R: *Measuring and assessing organisational culture in the NHS (OC1)*. London, UK: National Co-ordinating Centre for National Institute for Health Research Service Delivery and Organisation Programme (NCCSDO); 2008.
69. Weiner BJ, Amick H, Lee S-YD: **Conceptualization and measurement of organizational readiness for change: a review of the literature in health services research and other fields.** *Med Care Res Rev* 2008, **65**:379–436.
70. Schroder C, Medves J, Paterson M, Byrnes V, Chapman C, O'Riordan A, Pichora D, Kelly C: **Development and pilot testing of the collaborative practice assessment tool.** *J Interprof Care* 2011, **25**:189–195.
71. Avolio BJ, Bass BM, Jung DI: **Re-examining the components of transformational and transactional leadership using the Multifactor Leadership Questionnaire.** *J Occup Organ Psychol* 1999, **72**:441–462.
72. Irvine Doran DM, Baker GR, Murray M, Bohnen J, Zahn C, Sidani S, Carryer J: **Achieving clinical improvement: an interdisciplinary intervention.** *Health Care Manage Rev* 2002, **27**:42–56.
73. Wageman R, Hackman J, Lehman E: **Team Diagnostic Survey: Development of an Instrument.** *J Appl Behav Sci* 2005, **41**:373–398.
74. Hackman JR: **The design of work teams.** In *Handbook of organizational behavior*. Edited by Lorsch JW. Englewood Cliffs, NJ: Prentice-Hall; 1987:315–342.
75. Cohen SG, Bailey DE: **What makes teams work: Group effectiveness research from the shop floor to the executive suite.** *J Manage* 1997, **23**:239–290.
76. Campion MA, Medsker GJ, Higgs AC: **Relations between work group characteristics and effectiveness: Implications for designing effective work groups.** *Pers Psychol* 1993, **46**:823–850.
77. Campion MA, Papper EM, Medsker GJ: **Relations between work team characteristics and effectiveness: A replication and extension.** *Pers Psychol* 1996, **49**:429–452.
78. Shortell SM, Marsteller JA, Lin M, Pearson ML, Wu SY, Mendel P, Cretin S, Rosen M: **The role of perceived team effectiveness in improving chronic illness care.** *Med Care* 2004, **42**:1040–1048.
79. van den Bossche P, Gijssels WH, Segers M, Kirschner PA: **Social and Cognitive Factors Driving Teamwork in Collaborative Learning Environments: Team Learning Beliefs and Behaviors.** *Small Gr Res* 2006, **37**:490–521.
80. van der Vegt G, Emans B, van de Vliert E: **Motivating effects of task and outcome interdependence in work teams.** *Group Organ Manage* 1998, **23**:124–143.
81. Kirkman BL, Rosen B: **Beyond self-management: antecedents and consequences of team empowerment.** *Acad Manage J* 1999, **42**:58–74.
82. Langfred CW: **Autonomy and performance in teams: the multilevel moderating effect of task interdependence.** *J Manage* 2005, **31**:513–529.
83. Edmondson A: **Psychological safety and learning behavior in work teams.** *Adm Sci Q* 1999, **44**:350–383.
84. Alemi F, Safaie FK, Neuhauser D: **A survey of 92 quality improvement projects.** *Jt Comm J Qual Improv* 2001, **27**:619–632.
85. Lukas CV, Mohr DC, Meterko M: **Team effectiveness and organizational context in the implementation of a clinical innovation.** *Qual Manag Health Care* 2009, **18**:25–39.
86. Schouten L, Grol R, Hulscher M: **Factors influencing success in quality improvement collaboratives: development and psychometric testing of an instrument.** *Implement Sci* 2010, **5**:84.
87. Millward LJ, Jeffries N: **The team survey: a tool for health care team development.** *J Adv Nurs* 2001, **35**:276–287.
88. Bunderson J, Boumgarden P: **Structure and learning in self-managed teams: Why "bureaucratic" teams can be better learners.** *Organ Sci* 2010, **21**:609–624.
89. Thylefors I, Persson O, Hellstrom D: **Team types, perceived efficiency and team climate in Swedish cross-professional teamwork.** *J Interprof Care* 2005, **19**:102–114.
90. Lubomski LH, Marsteller JA, Hsu YJ, Goeschel CA, Holzmuller CG, Pronovost PJ: **The team checkup tool: evaluating QI team activities and giving feedback to senior leaders.** *Jt Comm J Qual Patient Saf* 2008, **34**:619–623. 561.
91. Manz CC, Sims HP: **Leading workers to lead themselves: The external leadership of self-managing work teams.** *Adm Sci Q* 1987, **32**:106–129.
92. Dobson RT, Stevenson K, Busch A, Scott DJ, Henry C, Wall PA: **A quality improvement activity to promote interprofessional collaboration among health professions students.** *Am J Pharm Educ* 2009, **73**:64.
93. Fulmer T, Hyer K, Flaherty E, Mezey M, Whitelaw N, Jacobs MO, Luchi R, Hansen JC, Evans DA, Cassel C, et al: **Geriatric interdisciplinary team training program: evaluation results.** *J Aging Health* 2005, **17**: 443–470.
94. Eby LT, Dobbins GH: **Collectivistic orientation in teams: An individual and group-level analysis.** *J Organ Behav* 1997, **18**:275–295.
95. Baker DP, Amodeo AM, Krokos KJ, Slonim A, Herrera H: **Assessing teamwork attitudes in healthcare: development of the TeamSTEPS teamwork attitudes questionnaire.** *Qual Saf Health Care* 2010, **19**:e49.
96. Heinemann G, Schmitt MH, Farrell M, Brallier S: **Development of an attitudes toward health care teams scale.** *Eval Health Prof* 1999, **22**:123–142.
97. Driskell JE, Salas E, Hughes S: **Collective orientation and team performance: development of an individual differences measure.** *Hum Factors* 2010, **52**:316–328.
98. Shaw JD, Duffy MK, Stark EM: **Interdependence and preference for group work: Main and congruence effects on the satisfaction and performance of group members.** *J Manage* 2000, **26**:259–279.
99. Jehn KA, Greer L, Levine S, Szulanski G: **The effects of conflict types, dimensions, and emergent states on group outcomes.** *Group Decis Negot* 2008, **17**:465–495.
100. Morgeson FP, DeRue DS, Karam EP: **Leadership in teams: A functional approach to understanding leadership structures and processes.** *J Manage* 2010, **36**:5–39.
101. Wilkens R, London M: **Relationships between climate, process, and performance in continuous quality improvement groups.** *J Vocat Behav* 2006, **69**:510–523.
102. Hoegl M, Gemuenden HG: **Teamwork quality and the success of innovative projects: A theoretical concept and empirical evidence.** *Organ Sci* 2001, **12**:435–449.
103. Thompson BM, Levine RE, Kennedy F, Naik AD, Foldes CA, Coverdale JH, Kelly PA, Parmelee D, Richards BF, Haidet P: **Evaluating the quality of learning-team processes in medical education: development and validation of a new measure.** *Acad Med* 2009, **84**:S124–127.
104. Hiller NJ, Day DV, Vance RJ: **Collective enactment of leadership roles and team effectiveness: A field study.** *Leadership Quart* 2006, **17**:387–397.
105. Anderson N, West M: **Measuring climate for work group innovation: Development and validation of the Team Climate Inventory.** *J Organ Behav* 1998, **19**:235–258.
106. Kuhn T, Poole MS: **Do conflict management styles affect group decision making? Evidence from a longitudinal field study.** *Hum Commun Res* 2000, **26**:558–590.
107. Schippers MC, Den Hartog DN, Koopman PL: **Reflexivity in teams: A measure and correlates.** *Appl Psychol Int Rev* 2007, **56**:189–211.
108. Brav A, Andersson K, Lantz A: **Group initiative and self-organizational activities in industrial work groups.** *Eur J Work Org Psychol* 2009, **18**:347–377.
109. de Jong B, Elfring T: **How does trust affect the performance of ongoing teams? The mediating role of reflexivity, monitoring, and effort.** *Acad Manage J* 2010, **53**:535–549.
110. Staples D, Webster J: **Exploring the effects of trust, task interdependence and virtualness on knowledge sharing in teams.** *Inform Syst J* 2008, **18**:617–640.
111. Janssen O, Van de Vliert E, Veenstra C: **How task and person conflict shape the role of positive interdependence in management teams.** *J Manage* 1999, **25**:117–142.
112. Tjosvold D, Wedley WC, Field RH: **Constructive controversy, the Vroom-Yetton model, and managerial decision-making.** *J Occup Behav* 1986, **7**:125–138.
113. Savelsbergh CM, van der Heijden BI, Poell RF: **The development and empirical validation of a multidimensional measurement instrument for team learning behaviors.** *Small Gr Res* 2009, **40**:578–607.
114. van den Bossche P, Gijssels W, Segers M, Woltjer G, Kirschner P: **Team learning: Building shared mental models.** *Instr Sci* 2011, **39**:283–301.

115. Mathieu JE, Gilson LL, Ruddy TM: **Empowerment and team effectiveness: an empirical test of an integrated model.** *J Appl Psychol* 2006, **91**:97–108.
116. Arnold JA, Arad S, Rhoades JA, Drasgow F: **The Empowering Leadership Questionnaire: the construction and validation of a new scale for measuring leader behaviors.** *J Organ Behav* 2000, **21**:249–269.
117. Anderson N, West M: *The team climate inventory: manual and users guide.* Windsor: Nelson Press; 1994.
118. Anderson N, West M: **The team climate inventory: The development of the TCI and its applications in team-building for innovativeness.** *Eur J Work Org Psychol* 1996, **5**:53–66.
119. Barczak G, Lask F, Mulki J: **Antecedents of team creativity: An examination of team emotional intelligence, team trust and collaborative culture.** *Creat Innov Man* 2010, **19**:332–345.
120. Caldwell DF, O'Reilly CA III: **The determinants of team-based innovation in organizations.** *The role of social influence. Small Gr Res* 2003, **34**:497–517.
121. Carless SA, De Paola C: **The measurement of cohesion in work teams.** *Small Gr Res* 2000, **31**:71–88.
122. Costa AC, Anderson N: **Measuring trust in teams: Development and validation of a multifaceted measure of formative and reflective indicators of team trust.** *Eur J Work Org Psychol* 2011, **20**:119–154.
123. Janssen O, Huang X: **Us and me: Team identification and individual differentiation as complementary drivers of team members' citizenship and creative behaviors.** *J Manage* 2008, **34**:69–88.
124. Bishop JW, Scott K, Goldsby MG, Cropanzano R: **A construct validity study of commitment and perceived support variables: A multifoci approach across different team environments.** *Group Organ Manage* 2005, **30**:153–180.
125. Bunderson JS, Sutcliffe KM: **Management team learning orientation and business unit performance.** *J Appl Psychol* 2003, **88**:552–560.
126. Lewis K: **Measuring transactive memory systems in the field: Scale development and validation.** *J Appl Psychol* 2003, **88**:587–604.
127. Irvine DM, Leatt P, Evans MG, Baker GR: **The behavioural outcomes of quality improvement teams: the role of team success and team identification.** *Health Serv Manage Res* 2000, **13**:78–89.
128. Shortell SM, Rousseau DM, Gillies RR, Devers KJ, Simons TL: **Organizational assessment in intensive care units (ICUs): construct development, reliability, and validity of the ICU nurse-physician questionnaire.** *Med Care* 1991, **29**:709–726.
129. Mokkink L, Terwee C, Knol D, Stratford P, Alonso J, Patrick D, Bouter L, De Vet H: **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.
130. Streiner DL, Norman GR: *Health measurement scales: a practical guide to their development and use.* 3rd edition. Oxford; New York: Oxford University Press; 2003.
131. Podsakoff P, MacKenzie S, Lee J: **Common method biases in behavioural research: a critical review of the literature and recommended remedies.** *J Appl Psychol* 2003, **88**:879–903.
132. van Ginkel WP, van Knippenberg D: **Group information elaboration and group decision making: The role of shared task representations.** *Organ Behav Hum Decis Process* 2008, **105**:82–97.
133. O'Toole TP, Cabral R, Blumen JM, Blake DA: **Building high functioning clinical teams through quality improvement initiatives.** *Qual Prim Care* 2011, **19**:13–22.
134. Cook JV, Dickinson HO, Eccles MP: **Response rates in postal surveys of healthcare professionals between 1996 and 2005: an observational study.** *BMC Health Serv Res* 2009, **9**:160.
135. Edwards PJ, Roberts I, Clarke MJ, Diguiseppi C, Wentz R, Kwan I, Cooper R, Felix LM, Pratap S: **Methods to increase response to postal and electronic questionnaires.** *Cochrane Database Syst Rev* 2009, MR000008.

doi:10.1186/1748-5908-8-20

Cite this article as: Brennan *et al.*: Measuring team factors thought to influence the success of quality improvement in primary care: a systematic review of instruments. *Implementation Science* 2013 **8**:20.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

