Psychometric properties of leadership scales for health professionals: a systematic review

Melissa A. Carlson1,2, Sarah Morris1,2, Fiona Day1,2,3, Ann Dadich4, Annika Ryan1,2, Elizabeth A. Fradgley1,2 and Christine Paul1,2*

Abstract

Background: The important role of leaders in the translation of health research is acknowledged in the implementation science literature. However, the accurate measurement of leadership traits and behaviours in health professionals has not been directly addressed. This review aimed to identify whether scales which measure leadership traits and behaviours have been found to be reliable and valid for use with health professionals.

Methods: A systematic review was conducted. MEDLINE, EMBASE, PsycINFO, Cochrane, CINAHL, Scopus, ABI/INFORMIT and Business Source Ultimate were searched to identify publications which reported original research testing the reliability, validity or acceptability of a leadership-related scale with health professionals.

Results: Of 2814 records, a total of 39 studies met the inclusion criteria, from which 33 scales were identified as having undergone some form of psychometric testing with health professionals. The most commonly used was the Implementation Leadership Scale (n = 5) and the Multifactor Leadership Questionnaire (n = 3). Of the 33 scales, the majority of scales were validated in English speaking countries including the USA (n = 15) and Canada (n = 4), but also with some translations and use in Europe and Asia, predominantly with samples of nurses (n = 27) or allied health professionals (n = 10). Only two validation studies included physicians. Content validity and internal consistency were evident for most scales (n = 30 and 29, respectively). Only 20 of the 33 scales were found to satisfy the acceptable thresholds for good construct validity. Very limited testing occurred in relation to test-re-test reliability, responsiveness, acceptability, cross-cultural revalidation, convergent validity, discriminant validity and criterion validity.

Conclusions: Seven scales may be sufficiently sound to be used with professionals, primarily with nurses. There is an absence of validation of leadership scales with regard to physicians. Given that physicians, along with nurses and allied health professionals have a leadership role in driving the implementation of evidence-based healthcare, this constitutes a clear gap in the psychometric testing of leadership scales for use in healthcare implementation research and practice.
The challenge of improving research translation or implementation

Translation of scientific knowledge to routine, evidence-based practice in healthcare settings ensures optimal care and improved outcomes for patients [1, 2]. Despite this, the translation of research knowledge to evidence-based practice is often slow or poor [3–6]. A foundational study by McGlynn [7, 8] found that during a two-year period between 1998 and 2000, patients in the United States receive 55% of evidence-based care with great variance in the rate of evidence-based care received among medical conditions. Furthermore, a 2005 systematic review by Schuster et al. [9] found 30–40% of patients were missing out on treatment that has been proven to be effective, while 20–25% of patients were receiving treatments that they do not need or that can cause them harm. A more recent Australian study by Runciman et al. [10] in 2012 with a sample of 1154 participants found that participants received appropriate care at 57% of healthcare encounters, again varying across medical conditions (from 32 to 86%). McGlynn [8] suggests that despite attempts to address these deficits in evidence-based care, there have been no large-scale studies in the United States measuring the provision of evidence-based care since 2003 and that although smaller studies indicate there have been improvements in some areas, there has been little change in healthcare overall. This failure to translate knowledge to evidence-based practice can result in poor outcomes for patients including sub-optimal treatment, exposure to unnecessary or harmful treatment, poorer quality of life, and loss of productivity [2, 6]. For healthcare systems, this failure can result in ineffective organisations and unnecessary expenditure [2, 6].

In healthcare, evidence-based practice refers to the translation or implementation of clinical research and knowledge into healthcare practice [6]. The two key steps toward evidence-based practice are: first, the translation of basic scientific knowledge to clinical practice, and secondly, the implementation of evidence-based practices that have found to be effective in the local setting into routine healthcare and policy [6, 11]. Barriers to successful implementation can be individual, structural, and organisational cultural [6, 12], including commitment from management, access to research, capacity issues, financial disincentives, inadequate skills within an organisation, or a lack of requisite facilities or equipment, staffing, peer morale and commitment, and leadership [6, 12]. Implementation strategies and frameworks assume or include important roles for leaders. Leadership has been shown to be an integral factor in nurturing a culture of evidence-based practice in clinical settings including cancer care, substance abuse, weight management, palliative care, and physiotherapy [3, 13–18]. Subsequently, leadership behaviours can encourage or discourage change and innovation within healthcare organisations [13, 19].

Despite leadership being considered a determining factor in implementing and sustaining evidence-based practices [1, 4, 20–24], the term remains an ambiguous concept in research [16]. Leadership has been conceptualised as a series of inherent personal traits, as learned behaviours, and as responses to particular situations or contexts [23]. Various types of leadership have been proposed including transformational leadership, transactional leadership, distributive leadership, charismatic leadership, heroic leadership, empowering leadership, engaging leadership, authentic leadership, collective leadership, servant leadership and passive or avoidant leadership [25–29]. A systematic review by Reichenpfader et al. [16] found that in 17 studies in the field of implementation science, the term was used imprecisely and inconsistently [16]. For the purpose of this paper, the authors will use Reichenpfader.
et al.’s [16] definition of leadership, being “a process of exerting intentional influence by one person over another person or group in order to achieve a certain outcome in a group or organization”. Likewise, the authors will consider leaders to be those people who are considered to exert influence on group or organisational outcomes, be they formal or informal leaders.

Formal leaders or positional leaders - managers or supervisors whose responsibilities include the oversight of staff, budgets, and operations - have the ability to procure supervisors whose responsibilities include the oversight of staff, budgets, and operations - have the ability to procure and disperse funding and resources, and design and enforce implementation policies [19, 30]. Formal leaders have the responsibility to ensure that healthcare organisations support the implementation of evidence-based practice through adequate funding and resources, supportive plans, practices, and strategies, as well as providing a work environment conducive to implementation [19]. The Consolidated Framework for Implementation Research (CFIR) [31] considers formal leaders to be the people who project manage and coordinate implementation. In healthcare settings the implementation of practice change in health care often requires leadership from multiple professional groups including nurses, physicians and allied health [32]. Powell et al. (2015) have suggested implementation strategies that leverage formal leaders including recruiting, designating, and training leaders for the change [33].

However, it is not only formal leaders who influence implementation. Change champions, who may be formal or informal leaders and are also referred to as opinion leaders, implementation leaders, facilitators, and change agents throughout the literature [34], also play a critical role in effective implementation [3, 19, 30]. Change champions are people within an organisation who are invested in implementing change, work hard to bring that change to fruition, are often personable, and are influential [3, 34]. Change champions may be frontline staff who may or may not have a formal management role, who frequently positively influence others’ attitudes or behaviours [3, 6, 30, 34]. Change champions acquire their influence through their demonstration of technical competence and accessibility and availability to their peers [6]. The CFIR suggests formal or informal change champions in implementation are those who are dedicated to supporting and driving implementation and influence attitudes toward implementation [31]. Implementation strategies utilising change champions identified by Powell et al. [33] include: identifying change champions, preparing them for the intervention and ensuring they are informed so they may influence the support of their colleagues [33]. It is these champions who have the responsibility to facilitate healthcare organisation climates being implementation-friendly through gaining support from senior management, formal leaders, as well as their peers [19].

Despite the critical role of both formal and informal leaders in facilitating the implementation of evidence-based practice in healthcare organisations, there is relatively little empirical study of how various aspects of leadership may be directly related to the efficacy or speed of research translation, or to the delivery of evidence-based practice [2]. Although it is clear that leadership is critical in the successful implementation and the sustainability of innovations [1, 35], it is unclear how the leadership traits and behaviours can be identified, measured, and developed [2, 3, 5, 19].

Consequently, the study of the relationship between leadership and research translation in healthcare requires accurate and relevant leadership scales. Leadership and change management is a growing area of scholarship [36–41], and some progress has been made on beginning to identify and synthesise scales which measure leadership traits and behaviours and to validate the psychometric properties of these scales [42–44]. Given the need for a variety of health professionals to be involved in the leadership of practice change; a leadership scale cannot be considered valid and reliable for administration with health professionals, until it is tested with a broad cross-section of such health professionals. However, a systematic review of general implementation scales (i.e. not leadership-specific) has highlighted a gap in the development and availability of validated scales which can be applied to the assessment of leadership traits and behaviours [45]. This gap inhibits the ability of implementation researchers and health professionals to identify evidence-based traits and behaviours which can facilitate identifying formal and informal leaders who may be integral in the promotion and delivery of evidence-based healthcare.

**Methods**

The aim of this systematic review was to identify published leadership scales that have psychometric properties (reliability, validity or acceptability) which have been assessed with clinical health professionals.

This review follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (see Additional File 1) [46]. The synthesis methods of this review were guided by Clinton-McHarg et al.’s [45] 2016 work which examined the psychometric properties of scales developed in public healthcare and community settings [45]. This review was registered with PROSPERO (Registration Number CRD42019121544).

**Search strategy**

MEDLINE, EMBASE, PsycINFO, Cochrane, CINAHL, Scopus, ABI/INFORMIT, and Business Source Ultimate were searched to identify relevant studies published in English between January 2000 and December 2018. A
A second search was conducted with the same criteria between January 2019 and January 2020. These time periods were selected to optimise currency of the findings and given very few (if any) relevant studies were published prior to 2000. Prior to the database searches being conducted, search terms were developed through and iterative process guided by the PICO (problem, population, intervention and comparison, and outcome) Statement [47, 48]. These terms were refined in consultation with a senior librarian from the University of Newcastle, Australia to capture the relevant studies and to ensure the correct use of Boolean operators, truncation, and subject headings. The selected search terms for all databases related to the key concepts explored, being healthcare leadership (problem), health clinicians (population), the type of scale (intervention and comparison), and assessment of psychometric properties (outcome), with additional terms related to health included for non-health focussed databases (population). The full search strategy for the MEDLINE database is shown in Fig. 1.

Eligibility
Publications were included if they: (1) were peer-reviewed journal articles reporting original research results; (2) reported data collected from or about practicing health professionals; (3) identified and assessed a leadership related scale for reliability, validity, or acceptability (See Table 1 for selection criteria and key definitions).

Study selection
The initial search yielded 4593 records. Of these, 1779 duplicate records were excluded. From the remaining pool of 2814 records, the titles and abstracts from a subset of 100 records that had been randomly selected were independently screened by two authors (CP and MC), to pilot the application of the inclusion and exclusion criteria. Title and abstracts from an additional subset of 500 randomly selected studies were then independently screened by the two authors (CP and MC) with the remaining screened by one author (MC). Studies that did not meet the inclusion criteria were excluded. The full-text manuscripts of the remaining 462 studies were then sourced. Of these 462 studies, the full text of 160 (~35%) studies were screened by two authors (MC and CP). The remaining 302 studies were screened by one author (MC). Of the 462 full text manuscripts screened, 274 did not meet the inclusion criteria and were subsequently excluded, leaving 188 eligible publications. After further discussion, the criteria for a leadership scale were refined to exclude any scales that did not specifically address leadership (i.e. those measuring burnout, implementation, non-technical skills, organisational context, patient safety, task/event-based leadership, or work roles). Using these criteria, a further 149 records were then excluded, leaving 39 records remaining for extraction (See Fig. 2 for PRISMA diagram).

Data collection process & data items
The following information was extracted and tabulated from publications that met the inclusion criteria: (1) author(s); publication year; setting (e.g., oncology, cardiology, etc.); country of study; participants (e.g., physicians, nurses, multidisciplinary, etc.); study aim; methods; leadership assessment (namely, type and name of scale or tool); outcome assessment; and findings. And (2) psychometric properties including face validity, content validity, internal reliability, test-retest reliability, construct validity, criterion validity, responsiveness, acceptability, feasibility, revalidation, cross-cultural validation, convergent validity, and discriminant validity.

Summary measures
Setting, sample, and characteristics of the innovation being assessed
Settings, sample, and characteristics of the innovation were extracted including the country and setting where the scale was validated, as well as the gender and profession of the sample and the sample response rate.
**Face and content validity**

Face validity assesses whether a scale is meaningful and relevant to those who use the scale [49]. Scales were considered to have face validity where administrators and/or test-takers agreed through a formal process that the scale measures what it is designed to measure [49]. Content validity assesses whether the scale fully captures the concept and sample it is designed to measure. The scale was considered to have content validity if the paper described how the items were selected and assessed, which revisions were made, and how they were made, or the theories and/or framework guiding the scale design [50].
**Internal reliability and test-retest reliability**
Scales or subscales were considered to have internal consistency if the Cronbach’s alpha was >0.70 [51]. Where a paper only reported a range of Cronbach’s alphas for the scale’s subscales and part of the range was <0.70, internal consistency was rejected. Repeated administration of a scale with the same sample and within 2–14 days was necessary to consider the scale’s test-retest reliability (i.e. a re-administration period outside of 2–14 days did not satisfy our criteria) [52]. Further, test-retest reliability was achieved if correlations between scores from the two administration time points had an intraclass correlation coefficient (ICC) of >0.70 [45, 50].

**Construct and criterion validity**
Exploratory and/or confirmatory factor analysis (EFA/CFA) results were primarily used to determine a scale’s construct validity (i.e. internal structure). If both an EFA and CFA were conducted for a single scale, cut-offs were applied to the CFA results. When interpreting an EFA, scales were considered to have construct validity if eigenvalues were set at >1 and/or >50% of variance was explained by the scale [53, 54]. In studies where percentage of variance explained was reported, eigenvalues of >1 were assumed. When interpreting a CFA, scales were considered to have construct validity where analysis was performed with a root mean square error of approximation (RMSEA) <0.08 and a comparative fit index (CFI) >0.95 [55, 56]. While a RMSEA of <0.06 is supported by Clinton-Mcharg (2016) [45], in this healthcare leadership literature, it was more common for an RMSEA of <0.08 to be an acceptable cut-off, as often referenced from Hu and Bentler (1999) [56]. A scale was considered to have criterion validity if different scores were obtained for subpopulations with known differences (e.g., general nurse versus nurse manager) [57].

**Responsiveness, acceptability, feasibility, revalidation, and cross-cultural adaptation**
A scale’s ability to detect change over time (i.e. responsiveness) was determined based on a reported moderate effect size (>5%) and/or minimal floor and/or ceiling effects (<5%) [50, 58]. A scale was considered acceptable based on a low proportion of missing items and feasible based on time taken to complete, interpret, and score the scale. It was also noted if a scale was revalidated with additional populations or samples, or adapted across cultures or languages.

**Convergent and discriminant validity**
A scale’s convergent and discriminant validity was determined respectively by Pearson’s correlation coefficients (r) >0.40 with similar scales and (r) <0.30 with dissimilar scales. Where convergent or discriminant validity was reported for a scale, however testing did not involve correlating the scale with other similar/dissimilar validated scales, these were marked as unclear when determining satisfaction of criteria.

**Synthesis of results**
Given that the publications varied considerably in their use and description of methodologies and measurements, a narrative synthesis rather than a meta-analysis was required. Popay et al. (2006:5) suggest that unlike a narrative review, which ‘are typically not systematic or transparent in their approach’, [59] narrative synthesis denotes ‘a process of synthesis that can be used in systematic reviews focusing on a wide range of questions, not only those relating to the effectiveness of a particular intervention ... [It] is part of a larger review process that includes a systematic approach to searching for and quality appraising research-based evidence as well as the synthesis of this evidence’. [59] For the purpose of this review, studies were synthesised according to their expressed aim(s).

**Results**
Of the 2814 records screened at the title and abstract stage, 2352 records were excluded. The 462 records remaining were screened at the full text stage. Of those records, 274 were excluded, leaving 188 eligible publications. After further discussion, the criteria for a leadership scale were refined to exclude any scales that did not specifically address leadership (i.e., those measuring burnout, implementation, non-technical skills, organisational context, patient safety, task/event-based leadership, or work roles). Using these criteria, a further 149 records were then excluded, leaving 39 unique records remaining for extraction (See Fig. 2 for PRISMA diagram).

**Study characteristics**

**Setting and characteristics of study sample for assessed scale**
Of the 33 scales, the majority of scales were validated in English speaking countries including the USA (n = 15) and Canada (n = 4), but also with some translations and use in Europe and Asia. The Implementation Leadership Scale was validated with five separate types of health professionals, more than any other of the included 33 scales. This was followed by the Multifactorial Leadership Questionnaire and the Evidence-Based Practice Nursing Leadership Scale, which were both validated in two separate types of health professionals. The majority of studies validated scales with nurses (n = 27), followed by allied health (n = 10), with only two studies validating scales with a sample that included physicians; and no scales were validated with most other types of health professionals. It is also worth noting that women were overwhelmingly represented within the sample. The percentage of women in the studies ranged from 39% to 99.5%, with the average...
percentage of women across the 26 studies that reported
gender being 75%. Given that the studies with the lowest
rates of women in their samples were those studies that
included non-nurse health professionals, this is likely due
to nursing being a female-dominated profession. These
data were reported in Table 2.

Psychometric properties of the scales including face and
content validity, internal reliability, test-retest reliability,
construct and criterion validity, responsiveness,
acceptability, feasibility, revalidation and cross-cultural
validation, were assessed and reported in Table 3.

Face and content validity
Of the 39 studies, face and content validity were evaluated
and satisfied in 18 and 33 studies (16 and 30 scales),
respectively.

Internal reliability
Of the included 33 scales, 29 scales (88%) achieved
internal consistency, as indicated by Cronbach’s alphas
$.70. All five studies reporting on the ILS indicated
data were reported in Table 2.

Acceptability, feasibility, revalidation and cross-cultural
acceptability, feasibility, revalidation and cross-cultural

Test-retest reliability
Of the 33 included scales, nine scales were tested for
test-retest reliability [62, 71, 80, 86, 88, 90–92]. Consid-
ering the Pearson’s correlation coefficient cut-off of $.70
alone, seven scales achieved adequate test-retest reliability
[62, 71, 80, 88, 90–92] and two did not [86, 87]. Re-
administration periods ranged from within 2–14 days
(n = 5) [71, 88, 90–92], between 14 and 30 days (n = 3)
[62, 80, 87], and one year [86]. Our criteria for adequate
test-retest reliability required both an r of $.70 and a re-
administration period of between 2 and 14 days. The five
scales re-tested within 2–14 days [71, 88, 90–92] fulfilled
this criterion. One scale [80] demonstrated high test-
retest reliability (r = .96) slightly outside the
recommended re-administration period (15 days post-
initial assessment), and was deemed successful in
satisfying our criteria.

Construct and criterion validity
Thirty-three studies reported their scale’s internal struc-
ture using either an EFA (n = 10) [includes PCA (n = 7)],
a CFA (n = 10), or both (n = 12). Of the five studies [19,
76–79] reporting on the ILS, three [19, 77, 78] reported
acceptable thresholds for good construct validity and two
[76, 79] did not. Of the remaining 26 scales, 54% (n = 14)
satisfied the acceptable thresholds for good construct
validity, in that the EFA indicated > 50% of variance ex-
plained by the final model and eigenvalues were set at > 1
and/or the CFA indicated acceptable RMSEA (< .08) and
CFI (> .95) values. Five scales were marked as marginally
unsuccessful (i.e. ‘N’*) [60, 74, 75, 87, 90] in satisfying our
criteria for construct validity, indicating either an RMSEA
value <.08 but not <.06, and/or a CFI value >.90 but not
>.95. One study [63] reported only the scale’s RMSEA
value (< .08) and so, was marked as unclear (‘U’) when
determining adequacy of construct validity (i.e. needing
both the RMSEA and CFI to determine adequacy). Two
further scales [82, 92] were marked ‘U’ as, although men-
tioning factor analysis or construct validity, they did not
report RMSEA or CFI values. Four scales [64, 67, 80, 95]
did not satisfy our criteria for adequate construct validity.

Of the 33 included scales, five scales [62, 68, 73, 75, 93]
demonstrated criterion validity and one [60] was marked
as unclear. Ten scales were correlated against existing
scales to evaluate convergent and/or discriminant validity,
as indicated by Pearson’s correlations (r). Eight of these
scales (including the ILS, as convergent validity was tested
and achieved in three of the five ILS studies) [60, 63, 66,
68, 74–76, 93] were considered to have convergent validity
(r > .40) and two scales (the iLead and the ILS) [19, 75, 79]
were considered to have both convergent and discrini-
mat validity (r < .30). Three studies [67, 76, 87] reported
on convergent and/or discriminant validity that did not
involve correlating the scales with other validated scales
and thus, were marked unclear (‘U’). Only one scale
(Survey of Transformational Leadership) [93] achieved
acceptable construct, criterion, and convergent validity.

Responsiveness, acceptability, feasibility, revalidation, and
cross-cultural adaptation
Of the 39 studies, only five reported on responsiveness,
three of which included scales that satisfied our criteria
for floor and ceiling effects of < 5% [62, 71, 90]. One scale
[67] had a small ceiling effect with scores skewed toward
the higher end of the scale (14–62% of people obtaining
the highest possible score for each item). The three papers
that reported on their scale’s acceptability [67, 90, 94] sat-
sified low proportions of missing items. Only one study
recorded the time taken to complete the scale (5–10 min)
[67]. Other studies mentioned the expected time to
complete the test in their methodology but did not record
actual time taken by test-takers. Of the eight scales that
| Scale name                                                                 | Measures                                                                 | Scale details                                                                 | Country | Sample size | Response rate | Gender of participants | Profession of participants |
|---------------------------------------------------------------------------|--------------------------------------------------------------------------|--------------------------------------------------------------------------------|---------|--------------|-----------------------|----------------------------|----------------------------|
| Aspiring leaders in Healthcare - Empowering individuals, Achieving excellence, Developing talents (AHEAD) [60] | Self-evaluated leadership competency in existing and emerging AHP leaders. | • 25 items • Likert (5-point) • 2 Domains: Values (inc. 9 subdomains; e.g., Commitment, Compassion, Integrity) and Skills (inc. 16 subdomains; e.g., Creativity; Talent Management; Writing) | Singapore | 106          | Not reported       | Females = 68%            | Allied Health Professionals |
| Authentic Leadership Inventory (ALI) [61]                                  | Employee’s perspectives on their managers’ authentic leadership characteristics. | • 14 items • Likert (5-point) • 5 Domains: Self-Awareness; Balanced Processing; Moral/Ethical Behaviour; and Relational Transparency | USA     | n = 85       | Not reported       | Not reported             | Nurses                      |
| Authentic Leadership Self-Assessment Questionnaire (ALSAQ-P) [62]         | Self-evaluated authentic leadership.                                      | • 16 items • Likert (5-point) • 4 Domains: Self-Awareness; Internalised Moral Perspective; Balanced Processing; Relational Transparency | Poland  | n = 3299     | Not reported       | Female = 98.5%            | Registered Nurses           |
| Authentic Nurse Leadership Questionnaire (ANLQ) [63]                      | Self-evaluated authentic leadership in nurse leaders.                     | • Item number not specified • response scale not specified • 5 Domains: Self-Awareness; Moral Ethical; Relational Integrity; Shared Decision Making; Caring | USA     | n = 309      | Not reported       | Female = 94%              | Registered Nurses           |
| Charismatic Leadership Socialised Scale [64]                              | Charismatic leadership in nurse managers.                                 | • 143 items • Likert (7-point) • Two parts: 1. Behaviour of the manager: Dimensions of the Charismatic and Instrumental Leadership 2. Behaviour of employees: Size of Commitment and Satisfaction, Motivation and Team Effectiveness | Brazil   | n = 211      | Not reported       | Female = 81.5%            | Registered Nurses           |
| Clinical Leadership Needs Analysis (CLeeNA) [65]                         | Self-reported leadership needs of nurses and midwives.                   | • 103 items • Likert (7-point) • 7 Domains: Self and Team Development; Staff and Care Delivery; Technology and Care Initiatives; Financial and Service Management; Leadership and Clinical Practice; Patient Safety and Risk Management; and Standards of Care | Ireland  | n = 321      | ~ 14.7%            | Female = 91%              | Registered Nurses           |
| Clinical Leadership Survey (CLS) [66]                                     | Self-evaluated clinical leader behaviours in staff nurses (basis in transformational leadership). | • 41 items • Likert (5-point) • 5 Domains: Challenging the Process; Inspiring a Shared Vision; Enabling Others to Act; Modelling the Way; and Encouraging the Heart | Canada   | n = 480      | 41%              | Female = 97%              | Registered Nurses           |
| Scale name                                                      | Measures                                                                 | Scale details                                                                                                                                                                                                 | Country | Sample size | Response rate | Gender of participants | Profession of participants |
|---------------------------------------------------------------|--------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|--------------|------------------------|---------------------------|-----------------------------|
| Clinician Safety Culture and Leadership Questionnaire [67]    | Clinician report of safety culture in hospitals at organisational and    | • Clinician report<br> • 35 items<br> • Likert (5-point)<br> • 3 Domains: Perceived Quality of Collaborating between Clinicians; Organisational Commitment to Patient Safety and Management Style; and Effectiveness of Healthcare Leaders in their Workplace | Australia | n = 1134     | Not reported          | Female = 39%             | Physicians<br> Allied Health Professionals |
|                                                              | departmental levels, across teamwork, safety climate and leadership in    |                                                                                                                          |         |               |                         |                           |                             |
|                                                              | healthcare.                                                              |                                                                                                                          |         |               |                         |                           |                             |
| Cotter Preceptor Selection Instrument (CPSI) [68]             | NPD practitioner perceptions of nurse attributes to assist nurse preceptor| • 14 items<br> • Likert (3-point)<br> • 10 Domains: Clinical Competence; Nursing Process; Transformational Leadership; Collaboration/Communication Skills; Professional Development; Conflict Resolution; Commitment; Flexibility; Empowerment; and Values | USA     | n = 13       | Not reported          | Not reported            | Nurses                      |
|                                                              | selection.                                                               |                                                                                                                          |         |               |                         |                           |                             |
| Director of Nursing Survey: Importance of Role Competencies [69]| DON's Perceived importance of DON role competencies.                      | • 28 items<br> • Likert (6-point)<br> • 5 Domains: Transformational Leadership Forces; Structural Empowerment Forces; Exemplary Professional Practice Forces; New Knowledge, Innovation, and Improvements Force; and Empirical Quality Outcomes Force | China   | n = 208      | 69%                    | Female = 99.5%          | Directors of Nursing       |
|                                                              |                                                                          |                                                                                                                          |         |               |                         |                           |                             |
| Evidence-Based Practice Nursing Leadership Scale [70]         | Staff nurse perceptions of support provided by unit level nurse managers  | • 10 items<br> • Response scale not specified<br> • Item examples: 'My manager provides time for me to engage in EBP' and 'My manager makes sure that I have access to relevant research on my unit' | USA     | n = 422      | 24%                    | Not reported            | Registered Nurses          |
|                                                              | to engage in EBP.                                                       |                                                                                                                          |         |               |                         |                           |                             |
| Evidence-Based Practice Nursing Leadership Scale (Chinese adaptation) [71]| (As above)                                                              | (As above)                                                                                                                                                                                                   | China   | n = 419      | 96%                    | Female = 96%              | Registered Nurses<br> Senior Nurses<br> Nurse Supervisors or above |
|                                                              |                                                                          |                                                                                                                          |         |               |                         |                           |                             |
| Healthcare Evaluation & Assessment of Leadership (HEAL) [72]  | Self-evaluated leadership competency based on the core principles of      | • 24 items<br> • Likert (5-point)<br> • 5 Domains: Critical Thinking; Emotional Intelligence; Teamwork; Selfless Service; and Integrity | USA     | n = 126      | 63%                    | Female = 50%            | Health professionals (internal medicine, paediatrics, surgery, radiology) |
|                                                              | patient centeredness.                                                   |                                                                                                                          |         |               |                         |                           |                             |
| Human Capital Competencies Inventory (HCIC) [73]              | Self-evaluated capital sustainability leadership attributes.              | • 16 items<br> • Likert (5-point)<br> • 4 Domains: Ethical Leadership; Sustainable Leadership; Mindful Leadership; Servant Leadership | USA     | n = 99       | 45%                    | Not reported            | Nurse Managers             |
| Scale name                                      | Measures                                                                 | Scale details                                                                 | Country    | Sample size | Response rate | Gender of participants | Profession of participants                                      |
|------------------------------------------------|---------------------------------------------------------------------------|--------------------------------------------------------------------------------|------------|-------------|----------------------|---------------------------------------------------------------|---------------------------------------------------------------|
| Human Capital Sustainability Leadership Scale (HCSLS) [74] | Self-evaluated capital sustainability leadership attributes.             | • 16 items • Likert (5-point) • 4 Domains: Ethical Leadership; Sustainable Leadership; Mindful Leadership; Servant Leadership | Italy      | EFA n = 207  | Not reported       | Female = 65%                                                  | Leaders from public and private health and care organisations |
| iLead [75]                                      | Employees’ perspectives on their managers’ active and passive implementation specific leadership attributes. | • 20 items • Likert (5-point) • 2 Domains: Active Leadership; and Passive Leadership • Active Leadership Subdomains: Exemplary Behaviours; Individualised consideration; Intellectual Stimulation; and Contingent Reward • Passive Leadership Subdomains: Passive Management-by-Exception’ and Laissez-Faire | Sweden     | n = 336     | 75% to study, of those 41% answered scale | Female = 90%                                                  | Health professionals (primary, psychiatric, rehabilitation, acute hospital care, and others) |
| Implementation Leadership Scale (ILS) [76]      | Clinician reports of strategic leadership attributes in supervisors, specific to EBP implementation. | • 12 items • Likert (5-point) • Domains: Proactive Leadership; Knowledgeable Leadership; Supportive Leadership; and Perseverant Leadership | USA        | n = 459     | 80%               | Female = 79%                                                  | Mental Health Clinicians                                      |
| Implementation Leadership Scale (ILS) [76]      | (As above)                                                                 | (As above)                                                                     | USA        | n = 323     | Not reported       | Female = 63%                                                  | Alcohol and Other Drug Treatment Professionals                |
| Implementation Leadership Scale (ILS) [77]      | (As above)                                                                 | (As above)                                                                     | USA        | n = 214     | 93%               | Female = 92%                                                  | Child welfare workers                                         |
| Implementation Leadership Scale (ILS) [78]      | (As above)                                                                 | (As above)                                                                     | USA        | S1 n = 200  S2 n = 284 | Not reported | Not reported |                                            | Registered Nurses                                             |
| Implementation Leadership Scale (ILS) [79]      | (As above)                                                                 | (As above)                                                                     | USA        | n = 136     | 88%               | Female = 76%                                                  | Mental health supervisors                                     |
| Kuopio University Hospital Transformational Leadership Scale (KUHTLS) [80] | Nurse staff reports of manager’s/unit director’s transformational leadership attributes. | • 47 items • Likert (5-point) • 7 Domains: Decision; Appreciation; Growth; Justice; Performance; Individuality; and Administration | Cyprus     | n = 315     | 79%               | Female = 49%                                                  | Nurses                                                       |
| Leadership and Management Inventory (LaMi) [81] | Self-evaluation by superiors and/or subordinates of leadership skills and abilities of first-line nurse managers. | • 56 items • Likert (5-point) • 6 Domains: Organisation and Assignment; Information and Communication; Social Skills and Interpersonal Relations; Development and Support of Subordinates; Co-operation; and Analysis and Decisions | Sweden     | S1 n = 149  S2 n = 197 | S1 = 50% S2 = 87.5% | Not reported | Registered Nurses Healthcare Personnel from unspecified professions |
| Scale name | Measures | Scale details | Country | Sample size | Response rate | Gender of participants | Profession of participants |
|------------|----------|---------------|---------|-------------|---------------|------------------------|---------------------------|
| Leadership Behaviour Description Questionnaire (Modified Version) [82] | Subordinate's preference for specific behaviours from their leaders. | • 15 items • Likert (5-point) • Leadership behaviours and organisational commitment | India | $n = 50$ | Not reported | Not reported | Paramedical Professionals |
| Leadership Competency Inventory (LCI) [83] | Self-reports from either individual staff members or managers on perceived degree of importance and development need for listed job competencies. | • 32 items • Likert (5-point) • 5 Domains: Personal Mastery; Managing Processes; Managing Resources; Leadership; and Managing Relationships | USA | $n = 323$ | Not reported | Not reported | Nursing homes and hospitals |
| Leadership Influence over Professional Practice Environments Scale [84] | Leaders' perceptions of their influence over the PPEs they oversee. | • 59 items • Likert (5-point) • 6 Domains: Collegial Administrative Approach; Internal Strategy and Resolve; Authority; Access to Resources; Leadership Expectations of Staff; and Status | USA | $n = 150$ | Not reported | Female = 70% | Health Professionals (Vice president, chief nurse, associate chief nurse) |
| Leadership Practices Inventory (LPI) [43] | Original LPI recommended when used as an educational tool. | Original LPI • 30 items • Likert (5-point) • 5 Domains: Challenging the Process; Inspiring a Shared Vision; Enabling Others to Act; Modelling the Way; and Encouraging the Heart | Canada | $n = 67$ | 61% | Not reported | Nurses |
| | Newly derived LPI recommended for use in nursing research to measure leadership practices of established and aspiring nurses. | Newly Derived LPI • 27 items • Likert (10-point) • 3 Domains: Cognitive; Behavioural; Supportive | |
| Multifactor Leadership Questionnaire [85] | Both leader self-evaluation and subordinate evaluation of their supervisor's transformational, active and passive leadership attributes. | • 78 items • Likert (5-point) • 9 Domains: Idealised Influence (Attributed); Idealised Influence (Behaviour); Inspiration Motivation; Intellectual Stimulation; Individualised consideration; Contingent Reward; Active Management-by-Exception; Passive Management-by-Exception; and Laissez-Faire Leadership | Canada | $n = 378$ | 38% | Female = 94% | Registered Nurses |
| Multifactor Leadership Questionnaire [86] | As above | (As above) | Finland | $n_1 = 423$ \(n_2 = 78\) | \(n_1 = 73\%\) | Female = 95% | Nurses |
| Nurse Leadership and Organisational Culture (N-LOC) [87] | Self-reported attributes of leadership style and organisational culture. | • 62 items • Likert (7-point) • Domains: Leadership Style; and Organisational Culture | Hong Kong | $n_1 = 295$ \(n_2 = 1146\) | \(n_1 = 63\%\) \(n_2 = 72\%\) | Female = 89\% \(n_2\) Female = 88.5\% | Nurses |
| Scale name                                      | Measures                                                                 | Scale details                                                                                      | Country       | Sample size | Response rate | Gender of participants | Profession of participants |
|------------------------------------------------|--------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------|---------------|--------------|----------------------|----------------------------|----------------------------|
| Quantum Leadership Scale [88]                  | Nursing administrators’ quantum skills, leadership characteristics and functions. | • 37 items<li>• Likert (5-point)<li>• 3 Domains: Quantum Skills; Quantum Leaders’ Characteristics; and Quantum Leaders’ Functions | Iran          | n = 25       | 100%                | Female & Married = 80% | Nurse administrators       |
| Questionnaire on Self Perception of Nurses of Exercise of Leadership and (QUEPTAEEL) Questionnaire on Perception of Nurse Technicians and LPNs of Exercise of Leadership (QUAPEEL) [89] | Perception of leadership practice and coaching processes from the perspective of leaders (nurses QUAPEEL) and followers (nurse technicians and LPNs; QUEPTAEEL). Both Scales: • 20 items<li>• Likert (6-point) and open- and closed-ended questions<li>• Knowledge about leadership; and questions regarding abilities/attitudes of leaders/followers in coaching leadership practice | Brazil        | n = 887      | Not reported | Not reported          | Nurses<br>Nurse Technicians |
| Spiritual Leadership Questionnaire [90]        | Nurses’ spiritual leadership attributes.                                  | • 35 items<li>• Likert (5-point)<li>• 8 Domains: Vision; Faith and Hope; Altruism; Inner Life; Calling; Membership; Organisational Commitment; and Productivity | Iran          | n = 400      | 91%                  | Female = 75.5%            | Nurses                     |
| Supportive Leadership Behaviour Questionnaire [91] | Supportive leadership behaviours.                                      | • 20 items<li>• Likert (5-point)<li>• 4 Domains: Support for Development; Integrity; Sincerity; and Recognition | Iran          | n = 731      | 94%                  | Not reported              | Head nurses<br>Nurses      |
| Supportive Supervisory Scale [92]             | Supervisory support and ability to develop and maintain positive relationships. | • 15 items<li>• Likert (5-point)<li>• 2 Domains: Respects Uniqueness and Being Reliable | Canada        | n = 22       | Not reported | Not reported              | Healthcare Aides           |
| Survey of Transformational Leadership [93]     | Approaches to the conceptualization and measurement of transformational practices | • 84 items<li>• Likert (5-point)<li>• 5 Domains: Idealised Influence; Intellectual Stimulation; Inspirational Motivation; Individualised Consideration; and Empowerment | USA           | n = 214      | Staff = 62% Leaders = 57% | Female = 63.5%            | Substance Use Workers      |
| Sustainment Leadership Scale (SLS) [94]        | Sustainment leadership attributes in first-level leaders and staff perceptions of leadership during sustainment. | • 12 items<li>• Likert (5-point)<li>• Four Domains: Proactive Leadership; Knowledgeable Leadership; Supportive Leadership; and Perseverant Leadership | USA           | n = 157      | 95%                  | Not reported              | Child Welfare Workers      |
| Unnamed leadership questionnaire for dental practitioners [95] | Good and poor leadership. | • 61 items<li>• Likert (7-point)<li>• Domains: Challenges the notion of leadership domains and uses two constructs Good Leadership and Poor Leadership | Turkey        | n = 237      | 23%                  | Female = 40%            | Dental Practitioners       |
Table 2 Characteristics of study sample for assessed scales (Continued)

| Scale name                                | Measures                                                                                     | Scale details                                                                 | Country | Sample size | Response rate | Gender of participants | Profession of participants |
|-------------------------------------------|---------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------|---------|--------------|-----------------------|-----------------------------|-----------------------------|
| Unnamed scale developed for study [96]    | Organisational culture, leadership behaviours and job satisfaction.                        | • 60 items<br>• Likert (5-point)<br>• Domains: Organisational Culture; Leadership Behaviour; Job Satisfaction | Taiwan  | n = 200      | 67%                 | Female = 99.5%            | Nurses                      |
Table 3  Summary of psychometric properties reported for each scale

| Title                                                                 | Face validity | Content validity | Construct validity | Criterion validity | Internal consistency | Test-retest reliability | Responsiveness | Acceptability/feasibility | Revalidation/cross-cultural | Convergent Discriminant |
|----------------------------------------------------------------------|---------------|-----------------|--------------------|--------------------|----------------------|-------------------------|----------------|---------------------------|-----------------------------|-------------------------|
| Aspiring leaders in Healthcare - Empowering individuals, Achieving excellence, Developing Talents (AHEAD) [60] (Allied Health Professionals) | –             | Y               | N*                 | U                  | Y                    | –                       | –              | –                         | Y                          | –                      |
| Authentic Leadership Inventory (ALI) [61] (Nurses)                   | –             | Y               | Y                  | –                  | Y                    | –                       | –              | –                         | –                          | –                      |
| Authentic Leadership Self-Assessment Questionnaire (Polish language version) (ALSAQ-P) [62] (Registered Nurses) | –             | Y               | Y                  | Y                  | Y                    | Y                       | –              | Y                         | –                          | –                      |
| Authentic Nurse Leadership Questionnaire (ANLQ) [63] (Registered Nurses) | –             | Y               | U                  | –                  | Y                    | –                       | –              | –                         | Y                          | –                      |
| Charismatic Leadership Socialised Scale (adaptation for Brazilian culture) [64] | Y             | Y               | N                  | –                  | N                    | –                       | –              | –                         | –                          | –                      |
| Clinical Leadership Needs Analysis Instrument (CEIeNA) [65] (Registered Nurses) | –             | Y               | Y                  | –                  | Y (only subscales reported) | –                       | –              | –                         | –                          | –                      |
| Clinical Leadership Survey (CLS) [66] (Registered Nurses)            | Y             | Y               | Y                  | –                  | N                    | –                       | –              | –                         | Y                          | –                      |
| Clinician Safety Culture and Leadership Questionnaire [67] (Physicians, Nurses, Allied Health Professionals) | –             | Y               | N                  | –                  | Y (only subscales reported) | –                       | N              | Y                         | U                          | U                      |
| Cotter Preceptor Selection Instrument (CPSI) [68] (Nurses)           | Y             | Y               | –                  | Y                  | Y (only subscales reported) | –                       | –              | –                         | Y                          | –                      |
| Director of Nursing Survey: Importance of Role Competencies [69] (Directors of Nursing) | Y             | Y               | –                  | –                  | Y (only subscales reported) | –                       | –              | –                         | Y                          | –                      |
| Evidence-Based Practice Nursing Leadership Scale [70] (Registered Nurses) | Y             | Y               | Y                  | –                  | Y                    | –                       | –              | –                         | –                          | –                      |
| The Evidence-Based Practice Nursing Leadership Scale (Chinese Adaptation) [71] (Registered Nurses, Senior Nurses, Nurse Supervisors or above) | Y             | Y               | Y                  | –                  | Y                    | Y                       | Y              | –                         | Y                          | –                      |
| Healthcare Evaluation & Assessment of Leadership (HEAL) [72] (Health professionals (internal medicine, paediatrics, surgery, radiology)) | –             | Y               | –                  | –                  | N                    | –                       | –              | –                         | –                          | –                      |
| Human Capital Competencies Inventory (HCCI) [73] (Nurse Managers)     | Y             | Y               | –                  | Y                  | Y                    | –                       | –              | –                         | –                          | –                      |
| Human Capital Sustainability Leadership Scale (HCSLS) [74] (Leaders from public and | –             | –               | N*                 | –                  | Y (only subscales reported) | –                       | –              | –                         | –                          | Y                      |
| Title                                                                 | Face validity | Content validity | Construct validity | Criterion validity | Internal consistency | Test-retest reliability | Responsiveness | Acceptability/ feasibility | Revalidation/ cross-cultural | Convergent | Discriminant |
|----------------------------------------------------------------------|---------------|------------------|--------------------|--------------------|----------------------|------------------------|----------------|-----------------------------|-----------------------------|------------|---------------|
| private health and care organisations                               | Y             | Y                | N*                 | Y                  | Y                    | –                      | –              | –                           | Y             | Y          | Y             |
| iLead [75] (Health professionals (primary, psychiatric, rehabilitation, acute hospital care, and others) | –              | –                | N                  | –                  | –                    | –                      | –              | –                           | Y             | U          | U             |
| Implementation Leadership Scale (ILS) [76] (Alcohol and Other Drug Treatment Professionals) | –              | –                | Y                  | –                  | –                    | –                      | –              | –                           | Y             | U          | U             |
| Implementation Leadership Scale (ILS) [77] (Child Welfare Workers)   | –              | –                | Y                  | –                  | –                    | –                      | –              | –                           | Y             | Y          | Y             |
| Implementation Leadership Scale (ILS) [78] (Registered Nurses)       | Y             | Y                | Y                  | –                  | –                    | –                      | –              | –                           | Y             | Y          | Y             |
| Implementation Leadership Scale (ILS) [19] (Mental Health Clinicians) | Y             | Y                | Y                  | –                  | –                    | –                      | –              | –                           | Y             | Y          | Y             |
| Kuopio University Hospital Transformational Leadership Scale (KUHTLS) [80] (Nurses) | –              | –                | N                  | –                  | –                    | –                      | –              | –                           | Y             | U          | U             |
| The Leadership and Management Inventory (LaMI) [81] (Registered Nurses, Healthcare Personnel) | –              | Y                | Y                  | –                  | –                    | –                      | –              | –                           | Y             | Y          | Y             |
| Leadership Behaviour Description Questionnaire (Modified Version) [82] (Paramedical Professionals) | –              | –                | U                  | –                  | –                    | –                      | –              | –                           | –             | –          | –             |
| Leadership Competency Inventory (LCI) [83] (Nursing Homes and Hospitals) | Y             | Y                | Y                  | –                  | –                    | –                      | –              | –                           | –             | –          | –             |
| The Leadership Influence Over Professional Practice Environments Scale [84] (Health professionals: Vice President, Chief Nurse, Associate Chief Nurse) | –              | Y                | Y                  | –                  | –                    | –                      | –              | –                           | –             | –          | –             |
| Leadership Practices Inventory (LPI) (3-factor) [43] (Nurses)        | –              | Y                | Y                  | –                  | –                    | –                      | –              | –                           | –             | –          | –             |
| Multifactor Leadership Questionnaire (MLQ) [85] (Registered Nurses)  | –              | Y                | Y                  | –                  | –                    | –                      | –              | –                           | –             | –          | –             |
| Multifactor Leadership Questionnaire (MLQ) (6-factor) [86] (Nurses)  | Y             | Y                | Y                  | –                  | –                    | –                      | –              | –                           | N             | –          | –             |
| Nurse Leadership and Organisational Culture (N-LOC) Questionnaire (2-factor) [87] (Nurses) | –              | Y                | N*                 | –                  | –                    | –                      | Y              | N                           | U             | U          | U             |
| Title                                                                 | Face validity | Content validity | Construct validity | Criterion validity | Internal consistency | Test-retest reliability | Responsiveness | Acceptability/feasibility | Revalidation/cross-cultural | Convergent validity | Discriminant validity |
|---------------------------------------------------------------------|---------------|------------------|--------------------|--------------------|---------------------|------------------------|----------------|---------------------------|-----------------------------|-------------------|------------------------|
| Quantum Leadership Scale [88] (Nurse Administrators)                | Y             | Y                | –                  | –                  | Y                   | Y                      | –             | –                         | –                          | –                 | –                      |
| Questionnaire on Self-Perception of Nurses of Exercise of Leadership (QUAPEEL) [89] | Y             | Y                | –                  | –                  | Y                   | –                      | –             | –                         | –                          | –                 | –                      |
| Spiritual Leadership Questionnaire (SLQ) [90] (Nurses)              | Y             | Y                | N*                 | –                  | Y                   | Y                      | Y             | Y                         | Y                          | –                 | –                      |
| Supportive Leadership Behaviours Questionnaire (SLB) [91] (Nurses, Head Nurses) | Y             | Y                | Y                  | –                  | Y                   | –                      | –             | Y                         | –                          | –                 | –                      |
| Supportive Supervisory Scale (SSS) [92] (Healthcare Aides)         | Y             | Y                | U                  | –                  | Y                   | Y                      | –             | –                         | –                          | –                 | –                      |
| Survey of Transformational Leadership (STL) [93] (Substance Use Workers) | Y             | Y                | Y                  | Y                  | Y                   | Y                      | –             | –                         | –                          | Y                 | –                      |
| Sustainment Leadership Scale (SLS) [94] (Child Welfare Workers)     | –             | Y                | Y                  | –                  | Y                   | Y                      | N             | Y                         | –                          | –                 | –                      |
| Unnamed Leadership Questionnaire for (Dental Practitioners) [95]    | –             | Y                | N                  | –                  | Y                   | Y                      | –             | –                         | –                          | –                 | –                      |
| Unnamed Scale Developed for the Study (Hospital Nurses) [96]        | –             | Y                | Y                  | –                  | Y                   | Y (only subscales reported) | –             | –                         | –                          | –                 | –                      |

Carlson et al. Implementation Science (2021) 16:85
underwent a process of revalidation in additional settings and subpopulations, five were successful in language retranslation and use with additional populations [62, 69, 71, 90, 91], two were unsuccessful within our criteria [64, 80] and one was unclear [87].

**Discussion**

The objective of the review was to inform healthcare implementation regarding appropriate scales for assessing traits and behaviours for identifying formal or informal leaders who can successfully implement change. Notably, a large number of scales (n = 33) were identified as having undergone some form of psychometric testing with health professionals. However, only three of the scales had been tested on multiple occasions. These were the Implementation Leadership Scale (n = 5), the Multifactor Leadership Scale (n = 2), and the Evidence-Based Practice Nursing Leadership Scale (n = 2). The implementation Leadership Scale was found to have sound: face validity and content validity with Registered Nurses; construct validity with Child Welfare Workers, Registered Nurses, and Mental Health Clinicians; internal consistency with Child Welfare Workers, Registered Nurses, and Mental Health Clinicians; convergent validity with Mental Health Supervisors and Mental Health Clinicians. The Multifactor Leadership Questionnaire was found to have acceptable face validity, content validity, construct validity, and internal consistency with Nurses. The Evidence-Based Practice Nursing Leadership Scale was found to have acceptable face validity, content validity, construct validity, and internal consistency with Nurses. The Supportive Leadership Behaviours (SLS), which has been validated with 638 workers in three Spanish speaking countries (Spain, Argentina and Mexico) [99], the Ethical Leadership Behaviour Scale (ELBS) [98], which has been validated with 405 workers in Brazil, the School Counsellors Leadership Survey (SCLS) [97], which has been validated with 776 school counsellors and school counselling supervisors in the USA, and the Implementation Leadership Scale (ILS) [77], which has been cross-validated with 214 child-welfare providers in the USA. Glasgow et al. [101] suggest that a scale with acceptable internal consistency may also have a high number of items and consequently be more burdensome for users [101]. They further suggest it may be more pragmatic to consider content validity [101], which assesses how well the scale measures the concept and sample it is designed to measure. Content validity was strong in most (n = 30) scales in this study, including the Implementation Leadership Scale, Multi-Factor Leadership Questionnaire and Evidence-Based Practice Nursing Leadership Scale.

The findings in relation to construct validity are potentially concerning in that only 15 of the 33 scales were found to satisfy the acceptable thresholds for good construct validity. This potential concern has not been clearly identified in the literature regarding testing of leadership scales with non-health professional samples [102–104]. For example, one study found that although a more recent revision of the Multifactor Leadership Questionnaire (MLQ) exhibited high internal consistency, previous literature employed older versions that lacked discriminant validity [102]. Another study testing the construct validity of the Servant Leadership Scale (SLS) found the construct validity to be sound, however, the authors suggested that previous studies had not adequately tested the construct validity of the scale [71].

In relation to the remaining psychometric characteristics – test re-test reliability, responsiveness, acceptability, cross-cultural revalidation, convergent validity, discriminant validity and criterion validity – very limited testing has occurred.

There are seven scales that stand out as likely to be psychometrically sound for use with health professionals (at least for nurses and allied health professionals), in that they are reported to have satisfied most of the reliability and validity criteria. Of the scales tested in the English-language, the iLead scale demonstrated good internal reliability and face, content, criterion, convergent and discriminant validity, and was only marginally outside our cut-off for having satisfied construct validity (CFI > .90 but not > .95). It is important to note that several studies decided to deem a CFI of > .90 as adequate for good construct validity. The Supportive Leadership Behaviours
Scale also satisfied internal and test-retest reliability, face, content, and construct validity, and was successfully revalidated. The Survey of Transformational Leadership (STL) demonstrated internal consistency and good construct, content, criterion, and convergent validity. Finally, the Implementation Leadership Scale has been evaluated several times and repeatedly demonstrates strong internal consistency, face and content validity, and convergent and discriminant validity. There are some inconsistencies in the scale’s construct validity, with two of the five evaluations of the ILS not satisfying our criteria for adequate construct validity. Of the scales tested in languages other than English, the Brazilian adaptation of the Charismatic Leadership Socialised Scale demonstrated inadequate construct validity and internal consistency, and so was not successfully revalidated. The Authentic Leadership Self-Assessment Questionnaire (Polish version) (ALSAQ-P) reported on and satisfied seven of the 11 criteria, including internal and test-retest reliability, content, construct and criterion validity, and evidence of good responsiveness and revalidation. The Persian version of the Spiritual Leadership Questionnaire (SLQ) demonstrated good internal and test-retest reliability and face and content validity. Moreover, the Persian SLQ was deemed responsive, acceptable and feasible, and achieved revalidation in Persian language. This scale, like the iLead scale, had a CFI of >.90 but did not meet our cut-off of a CFI > .95. The Chinese translation of the Evidence-Based Nursing Leadership Scale (EBP Nursing Leadership Scale) achieved internal and test-retest reliability and face and content validity. In summary, seven scales were found to have acceptable psychometric properties for use in healthcare, being the: Authentic Leadership Self-Assessment Questionnaire (Polish version), the iLead, the Spiritual Leadership Questionnaire (Persian version), the Supportive Leadership Behaviours Scale, the Evidence Based Nursing Scale (Chinese translation), and the Implementation Leadership Scale.

Few studies assessed the degree to which scale might be considered pragmatic, such as the time required to complete the scale or the acceptability and feasibility of the scale. Given the importance of identifying validated leadership scales in implementation science [45], and the key role of acceptability, feasibility, and cost (including time and resources) in assessing implementation outcomes [105], this represents a significant gap in the literature. However, it must be acknowledged that the search strategy did not focus extensively on pragmatic aspects of scales, for which tools are now emerging (e.g., Stanick, 2021) [106]. The availability of a quick, acceptable, and validated leadership scale would provide opportunities for researchers, leaders, and clinicians to assess health professionals in busy clinics for evidence-based leadership to drive evidence-based healthcare.

**Limitations**

Due to the diversity of the literature on leadership, the chosen set of search terms may have excluded some relevant studies. The review inclusion criteria resulted in the exclusion of a large number of studies relating leadership in the context of developing or demonstrating specific or technical skills (e.g., surgical skills). While these types of scales were considered too narrow or purpose-specific to be of benefit for assessing healthcare leadership more generally, it is possible that these scales could be potentially useful if adapted or modified. In addition, as noted by a number of authors [101], the pragmatic aspects of scales are important for implementation but have not been thoroughly addressed here. Inclusion of such assessment would be a useful addition to the field. The assessment of construct validity in this review focussed on factor analysis, as this was the approach generally taken in these studies. It is acknowledged that other approaches such as assessing a construct’s relation to theory are also important to establishing construct validity.

Additionally, women were overwhelmingly represented in the samples, perhaps due to the high number of scales validated with nurses. A working paper by the World Health Organisation (WHO) analysed gender equity in health professionals in 104 countries [107]. They found that women make up 67% of health professionals in the included countries, however in most countries, occupations such as physicians, dentists and pharmacists are mostly dominated by men, with professions such as nursing and midwifery mostly comprised of women [107]. A 2017 systematic review of medical leadership in hospital settings [108] found 28 studies exploring physician leadership. Of those 28 studies, nine found ‘leading change’ to be described as an activity performed by physician leaders. This suggests there may be a role for physicians as formal or informal change champions. Boateng et al. [109] propose that one component of best practice of scale development and validation is to do so with the population it is intended to be used with. Given that most of these scales have been validated primarily with nurses and allied health professionals who are predominantly female, it is difficult to claim that these scales are suitable for assessing leadership traits and behaviours in healthcare professional groups which are mostly male, or professional groups other than nurses and allied health professionals. Therefore, future work may consider validating these scales with a wider variety of health professionals.

**Conclusion**

There are seven scales which may be sufficiently sound to be used with nurses and allied health professionals. These are The Authentic Leadership Self-Assessment Questionnaire, the iLead scale, the Spiritual Leadership Questionnaire, the Supportive Leadership Behaviours...
Scale, The Survey of Transformational Leadership the Evidence-Based Nursing Leadership Scale and the Implementation Leadership Scale. There is a research gap in assessing leadership traits and behaviours of physicians and it appears that males have been underrepresented in some validation studies. Given the role of leadership in driving best practice in healthcare, there is a need for further psychometric assessment and validation of existing scales with physicians, males, and in assessing and understanding gender and cultural differences in implementation leadership. This serves to limit confidence with which the available scales can be used across health care disciplines in implementation research and practice, but also provides an opportunity for advancing the science of implementation leadership.

### Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13012-021-01141-z.

**Additional file 1. PRISMA 2009 Checklist.**

**Acknowledgements**

The authors would like to acknowledge Emma Sherwood’s contribution to conceptualisation and grant preparation for this study.

**Authors’ contributions**

AD, FD, AR, EF, and CP conceptualised this study. MC, SM, & CP coded and extracted the records and were major contributors in writing the manuscript. All authors read and approved the final manuscript.

**Authors’ information**

n/a

**Funding**

This study was funded by a Hunter Cancer Research Implementation Flagship Program.

**Availability of data and materials**

All data generated or analysed during this study are included in this published article and its supplementary information files.

### Declarations

**Ethics approval and consent to participate**

Not applicable.

**Consent for publication**

Not applicable.

**Competing interests**

The authors declare that they have no competing interests.

**Author details**

1Hunter Cancer Research Alliance, Newcastle, New South Wales, Australia.
2School of Medicine and Public Health, University of Newcastle, Callaghan, New South Wales, Australia.
3Calvary Mater Newcastle, Waratah, New South Wales, Australia.
4Centre for Oncology Education and Research Translation (CONCERT), Western Sydney University, Penrith, Australia.

---

**References**

1. Aarons GA, Farahnak LR, Erhart MG, Sklar M. Aligning leadership across systems and organizations to develop strategic climate to for evidence-based practice implementation. Ann Rev Public Health. 2014;35:255–74. https://doi.org/10.1146/annurev-publhealth-032013-182847.

2. Gifford WA, Holyoke F, Squires JE, Angus D, Brouwer L, Egan M, et al. Managerial leadership for research use in nursing and allied health care professions: a narrative synthesis protocol. Syst Rev. 2014;3(1):57. https://doi.org/10.1186/2046-4053-3-57.

3. Flodgren G, O’Brien MA, Parmelli E, Grimshaw JM. Local opinion leaders: effects on professional practice and healthcare outcomes. Cochrane Database Syst Rev. 2019(6):CD000125. https://doi.org/10.1002/14651858.CD000125.pub5.

4. Gifford W, Davies B, Tournageau A, Lefebre N. Developing team leadership to facilitate guideline utilization: planning and evaluating a 3-month intervention strategy. J Nurs Manag. 2011;19(1):121–32. https://doi.org/10.1111/j.1365-239x.2010.01401.x.

5. Gifford WA, Davies BL, Graham ID, Tousrangeau A, Woodend AK, Lefebre N. Developing leadership capacity for guideline use: a pilot cluster randomized control trial. Worldviews Evid-Based Nurs. 2013;10(1):51–65. https://doi.org/10.1111/j.1741-6787.2012.00254.x.

6. Grimshaw JM, Eccles MP, Lavis JN, Hill SJ, Squires JE. Knowledge translation of research findings. Implement Sci. 2012;7(1):1. https://doi.org/10.1186/1748-5908-7-50.

7. McGlynn EA, Asch SM, Adams J, Keesey J, Hicks J, DeCristofaro A, et al. The quality of health care delivered to adults in the United States. N Engl J Med. 2003;348(26):2635–45. https://doi.org/10.1056/NEJMsa022615.

8. McGlynn EA. Measuring and improving quality in the US: where are we today? J Am Board Fam Med. 2020;33(Supplement):S28–35.

9. Schuster MA, McGlynn EA, Brook RH. How good is the quality of health care in the United States? Milbank Q. 2005;83(4):643. https://doi.org/10.1111/j.1468-0016.2005.00403.x.

10. Runciman WB, Hunt TD, Hannaford NA, Hibbert PD, Westbrook JI, Coiera EW, et al. CareTrack: assessing the appropriateness of health care delivery in Australia. Med J Aust. 2012;197(2):100–5. https://doi.org/10.5694/mja12.10510.

11. Kitson A, Brook A, Harvey G, Jordan M, Marshall R, O’Shea R, et al. Using complexity and network concepts to inform healthcare translation. Int J Health Policy Manag. 2018;7(3):231. https://doi.org/10.15171/ijhpm.2017.79.

12. Geerligs L, Rankin NM, Shepherd HL, Butow P. Hospital-based interventions: a systematic review of staff-reported barriers and facilitators to implementation processes. Implement Sci. 2018;13(1):36. https://doi.org/10.1186/s13012-018-0726-9.

13. Choi M, Kim HS, Chung SK, Ahn MJ, Yoo JY, Park OS, et al. Evidence-based practice for pain management for cancer patients in an acute care setting. Int J Nurs Pract. 2014;20(1):60–9. https://doi.org/10.1111/ijn.12122.

14. Damschroder LJ, Hagedorn HJ. A guiding framework and approach for implementation research in substance use disorders treatment. Psychol Addict Behav. 2011;25(2):194. https://doi.org/10.1037/a0022284.

15. Damschroder LJ, Lowery JC. Evaluation of a large-scale weight management program using the consolidated framework for implementation research (CFIR). Implement Sci. 2013;8(1):51. https://doi.org/10.1186/1748-5908-8-51.

16. Reichenpfader U, Carlfjord S, Nilsen P. Leadership in evidence-based palliative care in nursing homes in Sweden: a qualitative study informed by the organisational readiness to change theory. Implement Sci. 2018;13(1):1. https://doi.org/10.1186/s13012-017-0699-0.

17. Dannapfel P, Nilsen P. Evidence-based psychotherapy culture: the influence of health care leaders in Sweden. Open J Leadersh. 2016;5(3):51–65. https://doi.org/10.5694/mja12.10510.

18. Aarons GA, Farahnak LR, Erhart MG, Sklar M. Aligning leadership across systems and organizations to develop strategic climate to for evidence-based practice implementation. Ann Rev Public Health. 2014;35:255–74. https://doi.org/10.1146/annurev-publhealth-032013-182847.

19. Aarons GA, Ehrhart MG, Farahnak LR. The implementation leadership scale (ILS): development of a brief measure of unit level implementation leadership. Implement Sci. 2014;9(1):45. https://doi.org/10.1186/1748-5908-9-45.

20. McGowen K. Physical exercise and cancer-related fatigue in hospitalized patients: role of the clinical nurse leader in implementation of interventions. Clin J Oncol Nurs. 2016;20(1):E20–E7. https://doi.org/10.1186/s13012-016-0242-z.
21. Long JC, Cunningham FC, Wiley J, Carswell P, Braithwaite J. Leadership in complex networks: the importance of network position and strategic action in a translational cancer research network. Implement Sci. 2013;8:122.

22. Harvey G, Kitson A. Translating evidence into healthcare policy and practice: single versus multi-faceted implementation strategies - is there a simple answer to a complex question? Int J Health Policy Manag. 2015;4(3):123–6. https://doi.org/10.1016/j.ijhpm.2015.04.

23. Health Workforce Australia. Leadership for the sustainability of the health system - part 1 - a literature review. Adelaide: Health Workforce Australia; 2012.

24. Health Workforce Australia. Health LEADS Australia: the Australian health leadership framework. Adelaide: Health Workforce Australia; 2013.

25. ZA K, Nawaz A, Khan I. Leadership theories and styles: a literature review. Leadership. 2016;16(1):1–7.

26. Fischer SA. Transformational leadership in nursing: a concept analysis. J Adv Nurs. 2016;72(11):2644–53.

27. Günzel-Jensen F, Jain AK, Kjeldsen AM. Distributed leadership in health care: the role of formal leadership styles and organizational efficacy. Leadership. 2016;14(1):10–33. https://doi.org/10.1177/1742715015644441.

28. Liu JT, Rowley C, Dinh CK, Qian D, Le HQ. Team creativity in public healthcare organizations: the roles of charismatic leadership, team job crafting, and collective public service motivation. Public Perform Manag Rev. 2019;42(6):1488–80. https://doi.org/10.1177/1091142119859067.

29. Harris J, Mayo P. Taking a case study approach to assessing alternative leadership models in health care. Br J Nurs. 2018;27(11):608–13. https://doi.org/10.12968/bjon.2018.27.11.608.

30. Gifford WF, Lewis KB, Eidh AC, Fiset V, Abd-Fatah T, Abeg AC, et al. Feasibility and usefulness of a leadership intervention to implement evidence-based falls prevention practices in residential care in Canada. Pilot Feasibility Stud. 2019;5:103. https://doi.org/10.1186/s40814-019-0485-7.

31. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci. 2009;4:52. https://doi.org/10.1186/1748-5908-4-52.

32. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci. 2009;4:52. https://doi.org/10.1186/1748-5908-4-52.

33. Davidson ES, Mitchell A, Beverly C, Brown LM, Rettingiani M, Walden M, et al. Psychometric properties of the authentic leadership inventory in the nursing profession. J Nurs Meas. 2018;26(2):364–72. https://doi.org/10.1080/10610274.2018.1450330.

34. Davé JS, Raval NN, Shaikh MS, Balasubramanian A, Atchamathy MJ. Development and validation of a leadership framework for public health and community settings and mapping of constructs against the consolidated framework for implementation research: a systematic review. Implement Sci. 2016;11(1):148. https://doi.org/10.1186/s13012-016-0512-5.

35. Clouten-McHarg T, Yoong SL, Tzelepis F, Regan T, Fielding A, Skelton E, et al. Psychometric properties of implementation measures for public health and community settings and mapping of constructs against the consolidated framework for implementation research: a systematic review. Implement Sci. 2016;11(1):148. https://doi.org/10.1186/s13012-016-0512-5.

36. Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M, et al. A review of methods for research synthesis in health policy and program evaluation. York: York Health economy Group; 2006.

37. Ribeiro Chavaglia SR, Dela Coleta MF, Dela Coleta JA, Costa Mendes IA, Ribeiro Chavaglia SR. Development and psychometric testing of the clinical leadership framework. Adelaide: Health Workforce Australia; 2013.

38. Pinto-Martins AR, Arndt R, Schaefer J, Da Silva FM, DeluibModal C, McShane R, et al. Psychometric properties of implementation measures for public health and community settings and mapping of constructs against the consolidated framework for implementation research: a systematic review. Implement Sci. 2016;11(1):148. https://doi.org/10.1186/s13012-016-0512-5.

39. Tourangeau AE, McGilton K. Measuring leadership practices of nurses using the leadership practices inventory. Nurs Res. 2004;53(3):182–9. https://doi.org/10.1097/00000519-200405000-00005.
needs an (CLeeNA) instrument for nurses and midwives. J Nurs Manag. 2019;27(2):245–55. https://doi.org/10.1111/jonm.12672.

66. Patrick A, Laschinger HK, Wong C, Finegan J. Developing and testing a new measure of staff nurse clinical leadership: the clinical leadership survey. J Nurs Manag. 2011;19(4):449–60. https://doi.org/10.1111/j.1365-2834.2011.01238.x.

67. Clay-Williams R, Taylor N, Ting HP, Winata T, Arnolda G, Braithwaite J. The clinician safety culture and leadership questionnaire: refinement and validation in Australian public hospitals. Int J Qual Health Care. 2020;32(Supplement_1),152–9. https://doi.org/10.1093/intqhc/mzz106.

68. Cotter E, Eckardt P, Moylan L. Instrument development and testing for selection of nursing preceptors. J. 2018;34(4):185–93. https://doi.org/10.11977/nnna.0000000000000464.

69. Spicer JG, Guo Y, Liu H, Hirsch J, Zhao H, Ma W, et al. Collaborative nursing leadership project in the People’s Republic of China. Int Nurs Rev. 2010;57(2):180–7. https://doi.org/10.1111/j.1466-7657.2009.0769x.

70. Phye Y, McDaniel A, Schafer J. Psychometric analysis of two new scales: the evidence-based practice nursing leadership and work environment scales. Worldviews Evid-Based Nurs. 2014;11(4):240–7. https://doi.org/10.1111/wvn.12045.

71. Zhang YP, Liu WH, Yan YT, Porr C, Zhang Y, Wei HH. Psychometric testing of the evidence-based practice nursing leadership scale and the work environment scale after cross-cultural adaptation in mainland China. Eval Health Prof. 2018;32(3):2781801439. https://doi.org/10.1177/01632871801801449.

72. Murphy KR, McManigle JE, Benjamin M, Wildman T, Jones AL, Dekker TJ, et al. Design, implementation, and demographic differences of HEAL: a self-guided educational leadership scale. J Healthc Leadersh. 2018:1042(2):233–32. https://doi.org/10.1111/j.1365-2834.2011.01238.x.

73. Donaher K, Russell G, Scoble KB, Chen J. The human capital competencies of first-line nurse managers. J Nurs Manag. 2008;16(7):784–9. https://doi.org/10.1111/j.1365-2834.2008.00912.x.

74. Acharya R, Dasbiswas AK. A study on the relationship between organizational commitment and leadership style on paramedical personnel in Kolkata. Int J Bus Insights Transformation. 2017;11(1):80–4.

75. Skytt B, Carlsson M, Ljunggren B, Engstrom M. Psychometric testing of the leadership and management inventory: a tool to measure the skills and abilities of first-line nurse managers. J Nurs Manag. 2008;16(7):784–94. https://doi.org/10.1111/j.1365-2834.2008.00912.x.

76. Ehrhart MG, Torres EM, Green AE, Trott EM, Willging CE, Moullin JC, et al. EBP implementation leadership scale (ILS) in substance use disorder treatment facilities. J Healthc Res. 2019;7:252–61. https://doi.org/10.302.https://doi.org/10.302.

77. Finn NK, Torres EM, Ehrhart MG, Roesch SC, Aarons GA. Cross-validation of the implementation leadership scale in acute care. Worldviews Evid-Based Nurs. 2019;16(4):141. https://doi.org/10.1111/wnn.12071.

78. Adams JM, Nikolaev N, Erickson JI, Ditomassi M, Jones DA. Identification of the psychometric properties of the leadership influence over professional practice environments scale. J Nurs Adm. 2013;43(5):258–65. https://doi.org/10.1097/NNA.0b013e31828eeb55.

79. Boamah SA, Tremplay P. Examining the factor structure of the MLQ transactional and transformational leadership dimensions in nursing context. West J Nurs Res. 2018;40(1):65–82. https://doi.org/10.11170/0136-2648.2006.04100.x.

80. Kanste O, Miettunen J, Kyrölä H. Psychometric properties of the multifactor leadership questionnaire among nurses. J Adv Nurs. 2007;57(2):201–12. https://doi.org/10.1111/j.1365-2640.2006.04100.x.

81. Luigi JINM, Johnston JM. Validation of the nurse leadership and organizational culture (N-LOC) questionnaire. BMC Health Serv Res. 2019;19(1):469. https://doi.org/10.1186/s12913-019-2430-z.

82. Dargahi H. Quantum leadership: the implication for Iranian nursing leaders. Acta Med Iran. 2013;51(6):411–7.

83. Cardoso ML, Ramos LH, D’Innocenzo M. Coaching leadership: leaders’ and followers’ perception assessment questionnaires in nursing. Einstein. 2014;12(1):66–74. https://doi.org/10.1590/1983-9750.05828.0408288.

84. Zagheri Tafreshi M, Jahanpard S, Rassouli M, Atashzadeh-Shosfideh F, Kavousi A. Psychometric properties of the Persian version of spiritual leadership questionnaire (SLQ): a methodological study. Iran Red Crescent Med. 2017;19(7):No pagination.e55930. https://doi.org/10.5812/icnmj.55930.

85. Shirazi M, Emani AH, Mirmoosavi SJ, Alavinia SM, Zamanian H, Fathollahbeigi F, et al. Contextualization and standardization of the supportive leadership behavior questionnaire based on socio-cognitive theory in Iran. Med J Islam Repub Iran. 2014;28:125.

86. McGilton KS. Development and psychometric testing of the supportive supervisory scale. J Nurs Scholarsh. 2010;42(2):223–32. https://doi.org/10.1111/j.1365-2834.2011.01238.x.

87. Edwards JR, Knight DK, Broome KM, Flynn PM. The development and validation of a transformational leadership survey for substance use treatment programs. Subst Use Misuse. 2010;45(9):1279–92. https://doi.org/10.1080/1082608100368038.

88. Ehrhart MG, Torres EM, Green AE, Trott EM, Willging CE, Moullin JC, et al. Leading for the long haul: a mixed-method evaluation of the sustainment leadership scale (SLS). Implementation Sci. 2018;13(1):17. https://doi.org/10.1186/s13012-018-0710-4.

89. Hill B, Brockehurst P. Leadership in dentistry: findings from new tool to measure clinical leadership. J Healthc Leadersh. 2015;7:13–20. https://doi.org/10.1111/j.1365-2834.2011.01238.x.

90. Tsai Y. Relationship between organizational culture, leadership behavior and job satisfaction. BMC Health Serv Res. 2011;11:98. https://doi.org/10.1186/1472-6963-11-98.

91. Young A, Bryan J. The school counselor leadership survey: instrument development and exploratory factor analysis. Prof Sch Couns. 2015;19(1):1–15. https://doi.org/10.1598/PSC.15030136.15030304.

92. Silva Filho ALA, Ferreira MC, Valentini F. Validity evidence of the ethical leadership behavior scale (ELBS). Psico-USF. 2019;24(2):349–59. https://doi.org/10.1590/1413-82712019240211.

93. Rodríguez-Carvajal R, de Rivas S, Herrero M, Moreno-Jiménez B, Van DierenDonck D. Leading people positively: cross-cultural validation of the servant leadership scale (GLS). Span J Psychol. 2014;17. https://doi.org/10.1017/spj.2014.73.

94. Cameli A, Reiter-Palmon R, Ziv E. Inclusive leadership and employee involvement in creative tasks in the workplace: the mediating role of psychological safety. Creat Res J. 2010;22(3):250–60. https://doi.org/10.808/08100419.2010.504654.

95. Glasgow RE, Riley WT. Pragmatic measures: what they are and why we need them. Am J Prev Med. 2013;45(2):237–43. https://doi.org/10.1016/j.amepre.2013.03.010.

96. Antounakis J, Avolio BJ, Sivasubramaniam N. Context and leadership: an examination of the nine-factor full-range leadership theory using the multifactor leadership questionnaire. Leader Q. 2003;13(4):261–95. https://doi.org/10.1016/S0149-984X(03)00304-4.

97. van Beversen P, Dimas ID, Lourenço PR, Rebelo T. Psychometric properties of the Portuguese version of the global transformational leadership (GTL) scale. Rev Psicol Trab Organ. 2017;33(2):109–14. https://doi.org/10.1186/s13012-018-0710-4.

98. Preacher KJ, Zhang Z, Zyphur MJ. Multilevel structural equation models for assessing moderation within and across levels of analysis. Psychol Methods. 2016;21(2):189. https://doi.org/10.1037/met0000052.
105. Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Health Ment Health Serv Res. 2011;38(2):65–76. https://doi.org/10.1007/s10488-010-0319-7.

106. Stanick CF, Halko HM, Nolen EA, Powell BJ, Dorsey CN, Mettert KD, et al. Pragmatic measures for implementation research: development of the psychometric and pragmatic evidence rating scale. Transl Behav Med. 2021;11(1):11–20. https://doi.org/10.1093/tbmn/lbz164.

107. Bonol M, McIsaac M, Xu L, Wuliji T, Diallo K, Campbell J. Gender equity in the health workforce: analysis of 104 countries: World Health Organization; 2019.

108. Berghout MA, Fabbricotti IN, Buljac-Samardžić M, Hilders CG. Medical leaders or masters?—a systematic review of medical leadership in hospital settings. PLoS One. 2017;12(9):e0184522. https://doi.org/10.1371/journal.pone.0184522.

109. Boateng GO, Neilands TB, Frongillo EA, Melgar-Quinonez HR, Young SL. Best practices for developing and validating scales for health, social, and behavioral research: a primer. Front Public Health. 2018;6:149. https://doi.org/10.3389/fpubh.2018.00149.

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