Quality indicators for dementia and older people nearing the end of life: A systematic review

Emel Yorganci MSc1 | Elizabeth L. Sampson MD2,3 | Juliet Gillam MSc1 | Jesutofunmi Aworinde MSc1 | Javiera Leniz MD1 | Lesley E. Williamson MSc1 | Rachel L. Cripps MSc1 | Robert Stewart MD4,5 | Katherine E. Sleeman PhD1

1Florence Nightingale Faculty of Nursing, Midwifery & Palliative Care, Cicely Saunders Institute, King’s College London, London, UK
2Marie Curie Palliative Care Research Department, Division of Psychiatry, University College London, London, UK
3Barnet Enfield and Haringey Mental Health Trust Liaison Psychiatry Team, North Middlesex University Hospital, London, UK
4Department of Psychological Medicine, Institute of Psychiatry, Psychology and Neuroscience, King’s College London, London, UK
5South London and Maudsley NHS Foundation Trust, London, UK

Abstract

Background: Robust quality indicators (QIs) are essential for monitoring and improving the quality of care and learning from good practice. We aimed to identify and assess QIs for the care of older people and people with dementia who are nearing the end of life and recommend QIs for use with routinely collected electronic data across care settings.

Methods: A systematic review was conducted, including five databases and reference chaining. Studies describing the development of QIs for care of older people and those with dementia nearing the end of life were included. QIs were categorized as relating to processes or outcomes, and mapped against six care domains. The psychometric properties (acceptability, evidence base, definition, feasibility, reliability, and validity) of each QI were assessed; QIs were categorized as robust, moderate, or poor.

Results: From 12,980 titles and abstracts screened, 37 papers and 976 QIs were included. Process and outcome QIs accounted for 780 (79.7%) and 196 (20.3%) of all QIs, respectively. Many of the QIs concerned physical aspects of care (n = 492, 50.4%), and very few concerned spiritual and cultural aspects of care (n = 19, 1.9%). Three hundred and fifteen (32.3%) QIs were robust and of those 220 were measurable using routinely collected electronic data. The final shortlist of 71 QIs came from seven studies.

Conclusions: Of the numerous QIs developed for care of older adults and those with dementia nearing the end of life, most had poor or moderate psychometric properties or were not designed for use with routinely collected electronic datasets. Infrastructure for data availability, combined with use of robust QIs, is important for enhancing understanding of care provided to this population, identifying unmet needs, and improving service provision.

KEYWORDS
dementia, end-of-life care, geriatrics, health care, quality indicators
BACKGROUND

The number of older people and people with dementia who have palliative care needs is increasing. The prevalence of dementia at death is around 30% in high-income countries, and with the aging population this number is rising globally. Delivering high quality care to meet complex needs arising from the interaction between physical and psychological needs, polypharmacy, clinical uncertainty, and care preferences can be challenging.

Monitoring and improving health care of populations require quality indicators (QIs), which can be obtained from routinely collected electronic datasets. Routinely collected electronic data can enable assessment and comparison of the care provided to people over time, across different care settings, nationally and internationally, while avoiding generating additional data collection tasks. QIs enable identification of services that deliver excellent care and those that might require improvement. At an aggregated level, QIs can also be used to assess the effectiveness of new interventions and services. QIs should ideally be accessible from routinely collected electronic datasets, supported by high-quality evidence and endorsed by key stakeholders.

Development and testing of QIs are resource-intensive and time-consuming processes. Numerous QIs have been developed to assess care of older people and those with dementia nearing the end of life. However, the properties and robustness of these QIs have not been systematically synthesized. To our knowledge, no systematic reviews have focused on QIs for the care of older people and those with dementia approaching the end of life. Although dementia is not synonymous with old age, there are similarities between the two patient cohorts toward the end of life in terms of demographics, and health and social care needs and service use.

Most people living with dementia are older than 60 years and people with dementia comprise 40%–70% of older adults living in long-term care facilities. Older people and people with dementia follow a similar, prolonged pattern of decline toward the end of life. The aims of this systematic review were: (1) to identify and assess the psychometric properties of QIs for the care of older people or people with dementia nearing the end of life and (2) to recommend QIs measurable using routinely collected electronic data across care settings.

METHODS

Search strategy

We searched MEDLINE, EMBASE, PsycINFO, Web of Science, and CINAHL from inception to February 2020. No study design or language limitations were applied. References of key studies and grey literature (i.e., publications produced by organizations outside the traditional academic publishing such as reports, working papers, and white papers) were searched. Search terms for each database are provided in Table S1. Definitions are provided in Table 1. Eligible studies (Table S2) were those describing the development, review, and/or testing of QIs for the care of adults with dementia and/or those who were older, and who were nearing the end of life.

When QIs were developed and evaluated over time and/or reported across multiple papers, data were extracted from the most recent publication, considering any refinements made. Publications reporting the application of existing QIs to clinical practice were used for reference chaining but excluded from the final list of papers. We excluded studies with adults younger than 60 years. Papers focusing predominantly on cancer and other disease-specific QIs (e.g., chronic obstructive pulmonary disease and end-stage renal failure) were also excluded. Service performance related QIs were excluded. Study authors were contacted to request additional information as required. The review was reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta analyses (PRISMA) reporting.

Key Points

- We identified 976 quality indicators (QIs); 508 QIs were judged to be robust, 231 moderate, and 237 poor in terms of their psychometric properties. The majority of the QIs were not designed for use with routinely collected electronic datasets.
- Few QIs have been developed regarding the legal, ethical, spiritual, and cultural aspects of care.
- 71 robust QIs for care of older people and people with dementia nearing the end of life, which can be used with routinely collected electronic data, across care settings are provided.

Why Does this Paper Matter?

The use of robust quality indicators with routinely collected electronic datasets will promote monitoring and improving the care provided to older people and people with dementia approaching the end of life.
**TABLE 1  Definitions**

<table>
<thead>
<tr>
<th>QIs</th>
<th>Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Routinely collected electronic data</td>
<td>Administrative data that were not predominantly collected for research purposes, those including electronic medical records, and data collected for insurance purposes such as mandated minimum datasets.</td>
</tr>
<tr>
<td>Population level</td>
<td>QIs which are applicable across the patient population, and not limited to a subpopulation or a care setting. While measures such as patient-reported outcome measures concern the quality of care provided at an individual level, QIs are meaningful measurements when they are interpreted for understanding the quality of care provided at an aggregated level.</td>
</tr>
</tbody>
</table>

Abbreviation: QI, quality indicator.

identified references were downloaded into EndNote reference manager for de-duplication. Title and abstract screening was performed by one reviewer (EY), and all papers were double-screened at the full-text stage (EY, JL, LT, and RLC). Papers with ambiguous content or with discrepancies regarding eligibility were discussed with a second reviewer until a consensus was reached.

Data from the included studies were extracted (EY, JG, and JA) into an Excel spreadsheet developed for the review (Link S4). Data extracted about each study included bibliographical information, aim, design, setting, country, population, and data source. Data extracted about each QI were based on a systematic review focusing on end-of-life cancer care QIs\(^{12}\) and the care domains adapted from the National Consensus Project’s guidelines for quality palliative care (developed in the United States and used by numerous organizations).\(^{23}\) Information retrieved included (1) QI type (process: what care is given and received; or outcome: changes in health status or quality of life\(^{21}\)); (2) care domain (“operational”; “physical”; “psychosocial”; “spiritual and cultural”; “communication, advance care planning, ethical and legal”; and “other [including QIs that cover multiple care domains]”;\(^{12,23}\); (3) numerator and denominator descriptions; and (4) information on psychometric properties (acceptability, evidence base, definition, feasibility, reliability, and validity). If the information required was not available in the included paper, backward and forward reference checks were made using Google Scholar.

**Data analysis**

The psychometric properties of each QI were double-checked and discussed between assessors (EY, JG, and JA) and other author (KES) where necessary. The six psychometric properties (acceptability, evidence base, definition, feasibility, reliability, and validity) have been shown to be important for assessment of individual QIs\(^{24-26}\) and similar criteria have been used previously for assessing sets of QIs.\(^{27}\) Assessment of psychometric properties was based on well-established criteria with four possible ratings: positive (positive supporting evidence [e.g., \( \geq 70\% \) of an expert QI development panel, for example, healthcare professionals and policymakers, agree that the QI is valid]), intermediate (doubts about the psychometric property [e.g., some but not all aspects of QI were clearly defined]), negative (disproving evidence about the assessed psychometric property [e.g., QI data collection, analysis, or interpretation not feasible]), and unknown (no published information available to make a judgment about the psychometric property [e.g., reliability of QI was not reported]).\(^{10,12,24}\) Further information on the definition and assessment of psychometric properties is provided in Table S5.

Rates of the psychometric properties were used to classify each QI’s overall usability as robust, moderate, or weak. QIs with three or more positive ratings across six psychometric categories were classified as robust. The final recommended shortlist included (1) QIs relevant across care settings (e.g., not just hospice), (2) QIs that were not reliant on a specific survey/data assessment tool, (3) robust QIs, and (4) QIs measurable using routinely collected electronic data. We used reference chaining to locate if QIs had been developed or tested for...
use with routinely collected electronic data such as electronic health records or insurance records. Where duplicate or related QIs were identified (e.g., depression screening for people diagnosed with dementia), the QI with (i) better psychometric properties and (ii) a broader denominator (greater sensitivity; applicable to more people) was chosen.28

RESULTS

Studies

We identified 12,980 publications from the electronic databases after de-duplication. An additional 19 publications were identified through reference chaining. Of these, 37 papers met the eligibility criteria (Figure 1) (Table S6). Studies were from the United States (n = 15),29–43 Canada (n = 6),44–49 The Netherlands (n = 3),50–52 Belgium (n = 3),53,54 the UK (n = 4),55–58 Australia (n = 2),59,60 Republic of Korea (n = 1),61 Japan (n = 1),62 Sweden (n = 1),63 Italy (n = 1),64 and Spain (n = 1).64 One paper covered QIs for Belgium, Italy, and Spain,64 and one paper conducted a comparative study between the United States, Europe, and Canada.65 Of 37 papers, 11 focused on care of older people nearing the end of life, nine on dementia end-of-life care, and 17 papers focused on end-of-life care of both populations. Eight papers described QIs focusing on care provided in hospital, 12 papers focused on community-based care, and 17 papers described QIs which could be applied across settings.

Quality indicators

From the 37 papers, 976 QIs (Link S4) were identified. Of these, 780 (79.7%) were categorized by type as “process” QIs and 196 (20.3%) as “outcome” QIs. The distribution of QIs by care domains was as follows: physical (n = 492, 50.4%); communication, advance care planning, and ethical and legal (n = 203, 20.8%); psychosocial (n = 117, 12.0%); operational (n = 110, 11.3%); multiple domains (n = 35, 3.6%); and spiritual and cultural (n = 19, 1.9%). Table 2 provides a summary of topics which were covered by the QIs within each care domain. Only 65 (6.9%) QIs were coupled with a benchmark value (i.e., a standard value against which the quality of care delivered can be measured). Just over half (n = 543, 55.6%) of QIs came from one QI set66 or adaptations of it to specific populations, care settings, or countries.32,33,38,48,62
Assessment of psychometric properties

Most of the QIs identified were rated as “positive” for at least one of the six psychometric properties (acceptability, evidence base, definition, feasibility, reliability, and validity) (Table 3). We were not able to make a judgment (marked as “unknown”) for at least one psychometric property in 846 (88.5%) QIs. Positive ratings of psychometric properties were as follows: acceptability (n = 786, 80.5%), evidence base (n = 760, 77.9%), definition (n = 512, 52.5%), feasibility (n = 494, 50.6%), reliability (n = 170, 17.4%), and validity (n = 614, 62.9%).

Overall, 508 QIs were judged to be robust, 231 moderate, and 237 weak in terms of their psychometric properties. Of the 508 robust QIs, 315 were unique QIs (i.e., not duplicated in other papers which developed QIs). When categorized into care domains, most of the unique, robust QIs focused on physical aspects of care (n = 204, 64.8%), followed by communication (n = 43, 13.7%), operational...
Two hundred and twenty robust QIs that could be used with routinely collected electronic data were identified. The shortlisted 71 QIs came from seven papers (Table S7). The majority (n = 54, 76.1%) of these QIs referred to a process of care, while the rest (n = 27, 23.9%) were outcome QIs. The final recommended QIs concerned physical (n = 29), operational (n = 17), communication, advance care planning and legal (n = 15), psychosocial (n = 9), and spiritual and cultural aspects of care (n = 1) (Figure 2). QIs concerning physical aspects of care varied from use of specific medications, appropriateness of medical interventions, and screening of symptoms and conditions. QIs on operational aspects of care mainly related to admissions, referrals, eligibility for specific care plans, or funding and

### TABLE 3 Summary of QIs by psychometric properties

<table>
<thead>
<tr>
<th>Psychometric property</th>
<th>QI rating N (%)</th>
<th>Positive</th>
<th>Intermediate</th>
<th>Negative</th>
<th>Unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acceptabilitya</td>
<td></td>
<td>786 (80.5)</td>
<td>71 (7.3)</td>
<td>3 (0.3)</td>
<td>116 (11.9)</td>
</tr>
<tr>
<td>Evidence baseb</td>
<td></td>
<td>760 (77.9)</td>
<td>83 (8.5)</td>
<td>1 (0.1)</td>
<td>132 (13.5)</td>
</tr>
<tr>
<td>Definitionc</td>
<td></td>
<td>512 (52.5)</td>
<td>329 (33.7)</td>
<td>116 (11.9)</td>
<td>19 (1.9)</td>
</tr>
<tr>
<td>Feasibilityd</td>
<td></td>
<td>494 (50.6)</td>
<td>NA</td>
<td>95 (9.7)</td>
<td>387 (39.7)</td>
</tr>
<tr>
<td>Reliabilitye</td>
<td></td>
<td>170 (17.4)</td>
<td>NA</td>
<td>15 (1.5)</td>
<td>791 (81.0)</td>
</tr>
<tr>
<td>Validityf</td>
<td></td>
<td>614 (62.9)</td>
<td>5 (0.5)</td>
<td>3 (0.3)</td>
<td>354 (36.3)</td>
</tr>
</tbody>
</table>

Abbreviation: QI, quality indicator.

aPerception among stakeholders that a QI is agreeable, or satisfactory measured within the degree of consensus.
bAvailability of scientific research and expert opinion regarding the process or outcome being measured as part of the QI.
cHow well a QI was defined.
dExtent to which a new QI can be successfully obtained (data collection) and analyzed.
eReliability measures (e.g., inter-rater, test-retest) relate to reproducibility of a QI.
fExtent to which a QI accurately reflects the domain of quality being assessed.

### FIGURE 2 Shortlist of recommended QIs for use with routinely collected electronic datasets (n = 71)

(n = 29, 9.2%), psychosocial (n = 28, 8.9%), other (n = 6, 1.9%), and spiritual (n = 5, 1.6%) aspects of care (Table 2). Two hundred and twenty robust QIs that could be used with routinely collected electronic data were identified.

**Shortlist of recommended QIs for use with routinely collected electronic datasets**

The shortlisted 71 QIs came from seven papers (Table S7). The majority (n = 54, 76.1%) of these QIs referred to a process of care, while the rest (n = 27, 23.9%) were outcome QIs. The final recommended QIs concerned physical (n = 29), operational (n = 17), communication, advance care planning and legal (n = 15), psychosocial (n = 9), and spiritual and cultural aspects of care (n = 1) (Figure 2). QIs concerning physical aspects of care varied from use of specific medications, appropriateness of medical interventions, and screening of symptoms and conditions. QIs on operational aspects of care mainly related to admissions, referrals, eligibility for specific care plans, or funding and

place of death. QIs in the communication, advance care planning, and legal category concerned communication of specific treatments to patients and their families and transfer of information between care settings. QIs relating to the psychosocial aspect of care were mainly regarding the screening and treatment of depression, people’s ability to communicate and socialize with others, and included three QIs related to carers. The single QI in the spiritual and cultural aspect of care concerned availability of translators (Table S7).

The shortlisted QIs included 35 (49.3%) which were applicable both to older people and people with dementia nearing the end of life, while 31 (43.7%) were for people with dementia and five (7.0%) were only developed for older people. Forty-three QIs were designed to be used for people nearing the end of life, while 28 QIs were designed to be used for people who had died (Figure 2).

**DISCUSSION**

**Main findings**

In this systematic review, we identified and assessed all available QIs for the care of older people and those with dementia nearing the end of life. We examined the psychometric properties of each QI, and their potential applicability at a population level using routinely collected electronic datasets. We produced a recommended shortlist of 71 QIs with robust psychometric properties which can be used for understanding, assessing, and improving care of older people and those with dementia nearing the end of life. Of the 71 shortlisted QIs, most concerned physical (n = 29) or operational (n = 17) aspects of care, while only one concerned spiritual or cultural aspects of care.

This distribution is similar to findings of systematic reviews of QIs for end-of-life care, cancer, or dementia care. Many of the shortlisted QIs in the physical care domain focused on multimorbidities, symptoms, and treatments, which reflect a biomedical approach to care of these populations. These QIs tend to be rated as more robust compared with the QIs in the other care domains. QIs about the operational aspects of care, such as those relating to admissions or eligibility for financial aid, were also common. Operational QIs such as hospital admissions before death are widely used by researchers and policy makers; hence, their psychometric properties may be better established. Information needed for understanding spiritual or cultural aspects of care is also less readily available in routinely collected electronic data, but may be captured in other ways such as through patient-reported outcomes and experience surveys. Thus, combining ways of measuring quality is key to gaining a broader overview of the quality of care.

Our shortlist also included a substantial number (n = 24) of QIs regarding communication; advance care planning; and ethical, legal, and psychosocial aspects of care. These QIs reflect receiving a diagnosis, discussion of treatments, communication of treatments to other professionals, receipt of treatment, and carers’ well-being. While application of the QIs in the communication, advance care planning, and ethical and legal domain may be less straightforward to extract from electronic systems, psychometric robustness of these QIs should encourage further adoption of these QIs into practice. We identified three QIs concerning carers available from routinely collected electronic data. Personal carers often have unidentified and unmet care needs. Using QIs related to carers is a promising starting point for recognizing their needs at a population level.

The smallest number of QIs (n = 19) concerned spiritual and cultural aspects of care. We identified one QI from this care domain for use with routinely collected electronic data, which referred to the availability of interpreters and translated materials. More QIs in the spiritual and cultural care domain have been developed in recent years. Their implementation into practice is currently limited. Spirituality and cultural beliefs are intrinsic aspects of person-centered end of life and palliative care and have been highlighted as one of the priority areas for palliative care of people with dementia. Building the evidence base and incorporating information regarding spiritual and cultural aspects of care in routinely collected national datasets may support the development of robust QIs in this area.

**Characteristics of QIs**

Over three quarters (79.7%) of the total identified QIs concerned processes of care, while 20.3% concerned outcomes. This may be because processes of care are often easier to measure and improve. Policy makers and health and social care providers can thus be more in control of setting benchmarks and making necessary changes to improve process QI measurements. Conversely, it is harder to determine the causal relationship between the care provided and the outcome experienced by people, which may explain the smaller number of outcome QIs.

In terms of psychometric properties, more than 50% of the QIs received a “positive” rating for their evidence base, acceptability, definition, and feasibility. Availability of information needed for assessing psychometric properties varied. Evidence base, acceptability, and definition of the QIs were well documented. In comparison, information...
on feasibility, reliability, and validity was harder to locate among the published work, reflected in higher percentages of “unknown” ratings for these psychometric properties. We also rated just 0.3% of the QIs as “negative” for their validity. Scant reporting of QIs with poor validity is likely to be explained by lack of validation of the QIs after their development. However, some level of publication bias may also exist, and some validation articles might have not been retrieved from our search.

While most QIs (n = 747, 76.5%) were applied prospectively to a population who were likely to be nearing the end of life, a smaller proportion (n = 229, 23.5%) of QIs were designed to be used retrospectively after death. Prospectively identifying a population approaching the end of life is challenging. QIs have been developed which signal potential palliative care needs or risk of poor outcomes such as functional and cognitive impairment and caregiver distress. However, these measures may not be available in routinely collected administrative data or health records and are not collected reliably across care settings. The mandate of standardized minimum datasets across settings may address this limitation in the future. Development of techniques such as linking different datasets and use of text-mining could also maximize the potential of obtaining QIs covering a wider range of care domains from routinely collected electronic data. Nevertheless, controls put in place for safety of patient data, organizational culture, complex governance arrangements including costs, technical barriers, and lack of transparency, and communication about routinely collected electronic data may create unintended barriers for its efficient use.

**Strengths and limitations**

The majority of systematic reviews for palliative care QIs have drawn on evidence from studies involving people with cancer. This is the first systematic review to identify and assess QIs applicable to older people and people with dementia nearing the end of life. We used previously applied and recommended ways of assessing psychometric properties of each individual QI rather than assessing the usability of overall QI sets. This approach enabled us to unpick QIs which were robust and usable within QI sets. All papers were double-screened at full-text stage, and the bibliographical data extracted and assessment of psychometric properties of each QI were also double-checked.

The evidence base for identified QIs was often unclear. In some cases, this was referred to as “based on existing evidence and expert opinion” without any further details. There is also a chance that information on the psychometric properties of some QIs has not been published. We made efforts by frequent discussions, reference chaining, and contacting the authors where needed, when making decisions on psychometric properties. QIs applicable across care settings have advantages such as comparison and continuity of measurements for a large population and are easier to apply. However, we acknowledge that setting- and condition-specific QIs have their own benefits. QIs, especially those focusing on processes and healthcare utilization, do not necessarily equate to achieving a good quality of care. Measures which may be relevant at a population level may not always translate to each individual’s experience and capture issues relevant to patients and families.

**Implications for policy, research, and practice**

Using QIs can help deliver high quality of care. The number of older people and those with dementia who are nearing the end of life is growing. With advancing technology, we can retrieve more information needed for evaluating palliative care and end-of-life care from routinely collected electronic datasets, while minimizing data collection burden. In light of the recent COVID-19 pandemic, the value of having linked datasets across care settings has been highlighted. Our findings should be used to inform the development of infrastructure needed for population-level data collection. Use of the recommended QIs may provide an overview of the quality of care provided to a large proportion of the population who are in significant need of palliative care, yet may consistently experience suboptimal care. Determining how many QIs to implement is difficult. Experts advise on being selective and having a smaller number of QIs per care domain. QIs are often pointers for actions needed to improve or maintain the quality of care. Therefore, having fewer robust QIs (rather than many collected through routinely collected electronic datasets) combined with other quality measures is more meaningful for achieving high quality of care.

Significant resource is spent for developing new QIs. We suggest that instead of developing new QIs, future research should prioritize the following: (1) revising and adapting existing QIs where possible (especially for physical aspects of care); (2) developing the evidence base for psychosocial, cultural, and spiritual aspects of care and further testing (feasibility, reliability, and validity) of existing QIs; (3) involving patients and families in further development and selection of QIs; (4) implementing shortlisted QIs into practice at local, national, and international population levels; and (5) combining the interpretation of the QIs with other quality measures such as patient-reported outcomes for monitoring and improving quality of care and enabling learning through comparison.
CONCLUSIONS

We provide a shortlist of 71 robust QIs for older people and people with dementia nearing the end of life that are robust, which can be applicable across care settings and measurable using routinely collected electronic datasets and applied across care settings. Future research should focus on testing and developing psychometric rigor of existing QIs and implementation of robust QIs into practice, to guide our understanding of quality of care provided to these populations, to identify unmet needs, and to improve service provision.

ACKNOWLEDGMENTS

We acknowledge Dr Lesley Henson for her advice in the initial stages of the study.

FUNDING INFORMATION

This project is funded by the National Institute for Health Research (NIHR) and Economic and Social Research Council (ESRC) ESRC/NIHR dementia initiative 2018 (Grant Reference Number ES/S010327/1). Katherine E. Sleeman is funded by an NIHR Clinician Scientist Fellowship (CS-2015-15-005), and is the Laing Galazka Chair in palliative care at King’s College London, funded by an endowment from Cicely Saunders International and the Kirby Laing Foundation. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

CONFLICT OF INTEREST

The authors have no conflicts.

AUTHOR CONTRIBUTIONS

Elizabeth L. Sampson, Katherine E. Sleeman, and Robert Stewart obtained funding for this study. Emel Yorganci, Katherine E. Sleeman, Elizabeth L. Sampson, and Robert Stewart designed the study. Emel Yorganci, Javiera Leniz, Lesley Thoms, Rachel L. Cripps, Jesutofunmi Aworinde, and Juliet Gillam were involved in data screening, extraction, and analysis. Emel Yorganci wrote the manuscript. All authors contributed to revisions, read, and approved the final manuscript.

SPONSOR’S ROLE

The funders played no role in the design, methods, subject recruitment, data collections, analysis, and preparation of paper.

ORCID

Emel Yorganci https://orcid.org/0000-0002-3903-1704
Elizabeth L. Sampson https://orcid.org/0000-0002-3692-8293
Katherine E. Sleeman https://orcid.org/0000-0002-9777-4373
Jesutofunmi Aworinde https://orcid.org/0000-0002-5977-6670
Javiera Leniz https://orcid.org/0000-0002-9315-4871
Lesley E. Williamson https://orcid.org/0000-0002-3807-7498
Rachel L. Cripps https://orcid.org/0000-0001-6945-5028
Robert Stewart https://orcid.org/0000-0002-4435-6397

REFERENCES


SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

Table S1. Search terms used in different databases.
Table S2. Eligibility criteria.
Table S3. PRISMA 2009 checklist.
Table S4. Data extraction sheet.
Table S5. Assessment of psychometric categories to evaluate quality indicators (QIs) which details rating criteria used to evaluate QIs.
Table S6. Summary of papers included in the systematic review.
Table S7. Summary of recurring topics covered in each quality indicator category.