

Protocol

Codeveloping a Set of Quality of Care Indicators for People Living With Multiple Long-Term Conditions in Primary Care: Protocol for a Modified RAND/UCLA Study

Sara Tavares¹, MSc; Thomas Beaney¹, PhD; Arad Reisberg², Prof Dr; Ania Henley²; Angela Edwards²; David Belsey²; Laura Downey^{3,4}, PhD

¹The George Institute for Global Health, Imperial College London, London, England, United Kingdom

²Patient Experience Research Centre (PERC), Imperial College London, London, United Kingdom

³The George Institute for Global Health, Imperial College London, Sydney, Australia

⁴Health Translation Hub, The George Institute for Global Health, UNSW Sydney, Sydney, New South Wales, Australia

Corresponding Author:

Sara Tavares, MSc
The George Institute for Global Health, Imperial College London
Scale Space, 58 Wood Ln
London, England W12 7RZ
United Kingdom
Phone: 44 07568150992
Email: s.tavares@imperial.ac.uk

Abstract

Background: Quality indicators in primary care remain predominantly disease-specific and professionally defined, with limited incorporation of what matters most to people living with multiple long-term conditions (MLTCs) and their caregivers. Existing frameworks and quality standards provide important conceptual direction, but few produce a pragmatic set of ready-to-use indicators.

Objective: This study aims to codevelop a disease-agnostic set of quality of care indicators with people living with MLTCs, caregivers, and health care professionals, including preliminary operational specifications to support future feasibility testing and service improvement in general practices within the National Health Service in England.

Methods: This protocol describes a 3-round modified RAND/UCLA Appropriateness Method (RAM) study. The first round will evaluate the importance of candidate indicators through an online questionnaire to an expert panel of patients, caregivers, and health care professionals. Indicators without agreement or requiring revision will proceed to a structured consensus meeting (round 2) with subsequent rerating rounds (round 3) for feasibility and measurability.

Results: As of July 2026, 10 patient and caregiver participants and 12 health care professionals have consented to join an expert panel. The expected RAM study completion, including data analysis and reporting, is early 2027. Fifty-four candidate quality indicators have been generated through triangulation of findings between a scoping review of published quality indicators in MLTCs and a qualitative study with 21 patients and caregivers with lived experience of MLTCs. Ongoing patient and public involvement and engagement with four community partners have further informed the study design, interpretation of multistage findings, and refinement of candidate indicators. The indicators have been mapped to a structure-process-outcome quality framework and will be taken forward into a modified 3-staged RAM study.

Conclusions: This study will generate a co-designed and implementation-oriented set of primary care quality indicators for MLTCs. The final indicator set is intended to support future measurement, quality improvement, and later field testing in real-world primary care systems.

International Registered Report Identifier (IRRID): DERR1-10.2196/99217

JMIR Res Protoc 2026;15:e99217; doi: [10.2196/99217](https://doi.org/10.2196/99217)

Keywords: quality indicators; primary care; multiple long-term conditions; multimorbidity; patient and public involvement; co-design; RAND/UCLA appropriateness method

Introduction

The coexistence of 2 or more long-term (chronic) conditions, commonly referred to as multiple long-term conditions (MLTCs) or multimorbidity, is a major global public-health problem with rates expected to rise over the next 30 years as populations live longer [1]. MLTCs affect populations across low, middle, and high-income countries [2], with global estimates suggesting that around 37.2% of adults living in the community are affected [3]. The prevalence of MLTCs increases with age, although the greatest number of people affected are those under 65 years old [3]. People with MLTCs are more likely to experience difficulties accessing and navigating health care services [4], functional limitations, poorer quality of life, reduced ability to work, and lower life expectancy [5]. MLTCs increase demand on health systems, particularly within primary care, where most receive their care [6]. People living with MLTCs also encounter unique challenges such as fragmented care with limited continuity between multiple health care providers, as well as poor coordination across services [7]. Subanalysis of patient experience surveys in primary care suggests that satisfaction with care declines as the number of conditions increases, and perceptions of care quality may differ from those without MLTCs [8].

To identify areas where the quality of care can be improved, collecting and interpreting health care performance data are crucial. Quality indicators, often referred to as quality measures or metrics, are one way to achieve this, by supporting the monitoring of the components of care associated with quality [9]. Rich evidence is available on the treatment burden for people with MLTCs [10,11], interventions to improve care [12,13], or individual components of care for chronic conditions [14,15]. However, limited consensus persists on which quality indicators best capture the quality of primary care delivered to people living with MLTCs [14]. Internationally published indicator sets for care quality in primary care often lack operational detail on definitions, numerators, denominators, and data sources, limiting direct transferability to current clinical care settings [16].

In the United Kingdom, while the National Institute for Health and Care Excellence (NICE) has produced a set of quality indicators intended to capture the processes of care relevant to those living with MLTCs, these have not been implemented through the national pay-for-performance program implemented in general practices (GPs) [17]. Through this program, payments to GPs are made according to the achievement of disease-specific quality indicator targets [18]. As a result, quality of care for people living with MLTCs continues to be monitored through aggregated analysis of these disease-specific Quality Outcomes Framework (QOF) indicators [18]. The limitation of these single-disease indicators is that they do not adequately reflect the complexity of living or caring for people with MLTCs, whose care and lived experience is not only shaped by their chronic conditions, but by an interplay of social, economic, and physical variables [19,20].

For care provider institutions, it is particularly challenging to measure holistic care that is relevant to patients and caregivers [21,22]. This is because a quality indicator needs to be able to measure a precisely defined activity for it to be reliable. Unlike clinical outcomes of care (eg, hospitalization or mortality), processes of care (eg, interpersonal relationships) are not consistently codified into standardized data fields within electronic health records and are often inserted within unstructured narratives, limiting their routine data extraction [23,24].

We aim to codevelop a set of quality indicators that better reflect the complexities of MLTC care, moving beyond the use of aggregated disease-specific metrics as measures of quality. Combining existing evidence, lived experience, and indicator development methodology via structured consensus, this study aims to codevelop a disease-agnostic set of quality-of-care indicators with people living with MLTCs, caregivers, and health care professionals, including definitions, preliminary numerators, denominators, exclusions, and data sources to support future feasibility testing and service improvement in primary care settings within the National Health Service in England.

Methods

Study Design

This study will adopt a modified RAND/UCLA Appropriateness Method (RAM) Delphi design, a structured consensus method widely used to develop and validate quality indicators across health care settings, including in psycho-oncological care [25], primary care for older adults [26], specialist mental health referral pathways [27], delivery of person-centered care [28], and various single-disease indicators [29-31]. By combining evidence from the literature with structured expert judgment, RAM supports the development of indicators that are evidence-informed, transparent, and reproducible, while strengthening face and content validity through consensus. In addition, its flexible format allows the feasibility of data collection questions, which is crucial to guide future clinical implementation strategies [9,32,33].

This RAM study will comprise multiple components: (1) a web-based first-round rating of candidate indicators; (2) a structured hybrid consensus meeting to discuss indicators not reaching agreement during round 1; and (3) a final rerating round to obtain final consensus on definitions, importance, and feasibility of data collection using routine primary care electronic records. When completed, the final report of this RAM study will follow the Delphi recommendations for an interdisciplinary standardized reporting (DELPHISTAR) [34].

Patient and Public Involvement

This study is underpinned by a pragmatic-interpretivist paradigm, which recognizes that knowledge intended to improve health care must be both grounded in multiple perspectives and lived experience while focusing on practical and applicable solutions for real-world practice

[35]. Pragmatism supports the use of multiple methods, data sources, and forms of expertise to address the research question and support future implementation [36,37], while interpretivism emphasizes understanding how the subjective experiences of patients and caregivers shape knowledge and perceptions of care quality in the context of everyday life [38].

This study will be explicitly grounded in principles of coproduction, collaboration, and inclusion, adopting an approach to conducting research *with* patients and caregivers rather than *for* them [39]. Their active participation during this RAM study will ensure that emerging indicators remain meaningful to those affected by MLTCs and relevant to clinical practice. Previous and future contributions from our community partners (CPs) are reported in detail following the Guidance for Reporting Involvement of Patients and the Public (GRIPP) framework [40] (Checklist 1) ().

Conceptual Framework Through a Scoping Review and Qualitative Interviews

The first step in indicator development requires a conceptual framework and rigorous methods to identify, clarify, and define what concepts and dimensions of quality should be measured [32,41,42]. An international scoping review is published elsewhere [43], and it synthesized evidence on the processes and content of existing MLTC-related quality indicators, including the extent to which patients had been involved in their development. Seventy-eight disease-agnostic indicators were extracted eliciting patient and/or caregiver priorities, and 33 quality domains were synthesized for future indicator development. Across studies, quality was more commonly articulated through processes of care than through organizational structures or outcomes, with recurring emphasis on coordination, continuity, shared decision-making, and holistic and tailored assessments. Few indicators were fully operationalized for routine primary care use, and many were measured through patient-experience surveys rather than clearly specified indicators with defined data sources or measurement rules.

Second, interviews with 21 adults living with or caring for someone with MLTCs explored how high-quality care is understood from lived experience. Through reflexive inductive thematic analysis, four interrelated domains of quality emerged: optimal clinical care, patient and carer centeredness, organizational processes and access, and system preparedness and support. This qualitative work showed that health care quality was dynamic, with priorities shifting over time according to symptoms, treatment burden, care transitions, and changing life circumstances. The full manuscript is under review elsewhere.

Panel Eligibility and Selection

Recommendations from the RAM suggest a panel size of between 7 and 15 members to allow a focused discussion, with the opportunity for different perspectives to be expressed [33]. We will aim to recruit approximately 15 panel members, with a planned stakeholder composition of

at least 7 patients and/or caregivers, 3 health care professionals from different professional roles, 1 commissioner, 1 service or operational manager, and 2 researchers or clinical academics. This planned composition is intended to ensure that patients, caregivers, and community advisory partners together represent at least half of the panel. Although there is no universal standard definition for who qualifies as an expert panel member in Delphi studies [44], we will consider the following as expert panelists: health care professionals (HCPs), commissioners, managers, or researchers with at least 5 years of experience relevant to primary care provision, quality improvement, or MLTC care. Publication history will not be an eligibility criterion because this study considers implementation and service insight alongside methodological or theoretical knowledge [45]. As such, we will include medical professionals, allied professionals, and nurses, if experience was established through clinical practice, organizational leadership, policy work, or academic expertise. The mixed sample of panelists background and expertise will account for a wide variation of relevant views to the care of those with MLTCs, avoiding homogeneous approaches to quality of care in MLTCs and mitigating cognitive biases [45].

All 21 patients and caregivers who participated in the earlier qualitative interviews will be invited to take part in the RAM panel, supporting the continuity of involvement across this multistep study, and to help reduce attrition between rounds. These participants have lived experience of MLTCs and receive care across primary and secondary care services. Many are also engaged in local community activities, patient-led charities, or research involvement initiatives. If necessary, the recruitment of additional patient and caregiver participants will be supplemented through the networks of our community advisory partner and Imperial's Patient Experience Research Centre, including engagement with local community organizations and patient-led voluntary groups. Patients and caregivers will be eligible if they are aged 18 years or older, have relevant self-reported lived experience of living with or caring for someone with MLTC, are able to provide informed consent and participate in the rating rounds and consensus meetings. Commissioners and service managers will be identified through repeated engagement with the local Integrated Care Board network. Health care professionals and clinical academics will be approached via purposive and snowballing approaches from the research team's academic and professional networks.

All participants will be compensated for their time with a high-street voucher at the completion of each round. To reduce attrition, we will provide clear information on expected time commitment for each round, schedule round 2 consensus meeting with at least 4 weeks in advance, send weekly reminders for the completion of surveys, and offer technical support. Between all rounds, HCP participants will receive data synthesis reports, while patient and caregiver participants will receive lay summaries of the findings.

Although we aim to retain a consistent core panel across this study's rounds to ensure that participants can consider group feedback and rerate the same items over time [46],

we will also allow new panel members who were unable to participate in round 1 to join the consensus meeting in round 2. Panelists who missed a previous round will continue to be invited to participate unless they have expressly withdrawn from the study. Boel et al [47] found that retaining these panelists does not significantly affect Delphi outcomes and may improve response rates. If additional or replacement panelists are recruited between rounds, recruitment will be purposive and monitored to preserve the balance of the panel. Where possible, participants will be recruited from the same stakeholder group (ie, pharmacists, nurses, or GPs) and will have broadly comparable professional background or lived experience of MLTCs to those who have withdrawn. Changes in panel composition across rounds will be reported transparently in the study flow diagram and participant characteristics table.

Preparing for Round 1: Web-Based Survey

The findings from both the scoping review and the qualitative interviews were triangulated and mapped onto a structure-process-outcome framework informed by the Donabedian model (Table S1 in Multimedia Appendix 1) [48]. This identified areas of convergence between published evidence and qualitative lived experience, identifying the main domains of care quality and the existing indicators developed in the literature.

To develop indicators, the RAM does not prescribe a specific syntactic format and suggests that items are

expressed as clinical scenarios or system attributes that can be judged for rating [33]. At this stage, 54 candidate indicators were drafted as clear declarative statements. Indicators following traditional “If-Then” definitions [9] were intentionally avoided to support a balanced assessment of validity, importance, and redundancy without eliciting premature operational framing.

Traditional RAM panels often use appropriateness as the main rating construct, defined by whether the expected health benefits of an intervention outweigh its potential harms [26, 33]. Because our indicators relate to structures and processes of primary care, rather than discrete single-disease interventions, during this round 1 all participants will be asked to vote on the importance of candidate indicators, rather than on appropriateness. Importance will be defined as the extent to which an indicator addresses an aspect of care encountered in routine primary care and is linked to care components that matter to patients, caregivers, quality improvement, and primary care clinical activities [32,42,49].

Guidance on rating quality indicators is extremely heterogeneous without a universally used set of questions for its assessment, but there is agreement that robust indicator development should at minimum address content and face validity (Table 1) [9,32,42,50]. Validity can be assessed using varying methods, and it usually asks patients or HCPs about the relevance, comprehensiveness, and comprehensibility of the items [33,49].

Table 1. Criteria for indicator rating and assessment.

Criteria for indicator rating and assessment	Description
Content validity	The extent to which the indicator accurately measures what it is intended to measure is underpinned by evidence and captures meaningful aspects of the quality of care, consistent with professional knowledge and the outcomes of high-quality care [32,42,49, 50].
Face validity	The extent to which the indicator makes sense logically and clinically [50], often reflected as to whether they link with structure, processes, and/or outcomes in the health care system [32]. Campbell et al [9] defines it as whether any indicators were rated by panels and if there is evidence that the indicator is underpinned by consensus.
Relevance, importance, or acceptability	These terms are usually used interchangeably in published studies developing indicators. Relevance can describe either the impact of disease on health expenditure and organizational contexts. Also relates to the stakeholders’ views on its importance as a quality gap or potential for performance improvement as a gap in performance [27,51]. Acceptability describes the extent to which indicator contents are acceptable to those being assessed and those undertaking the assessment. It is primarily used during application of quality indicators, rather than development stages [9].
Redundancy	Although many methodological recommendations focus on single indicators, indicator sets require explicit attention to omission, redundancy, overrepresentation of some domains, inclusion of irrelevant indicators [42], or existing related and competing measures already in practice [52].

Patients and caregivers will rate the importance based on their lived experience of MLTCs, while HCPs will additionally rate the importance of indicators for measuring overall quality of care and to primary care clinical activities. Ordinal ratings will be used following a 9-point Likert scale, aligning with RAM recommendations facilitating granular measurement [33,53].

Advised by our CPs, patients and caregivers will not be asked to vote importance to quality improvement or primary care clinical activities, as these constructs were considered less accessible and relevant from lived experience alone. Table 2 summarizes round 1 rating assessment criteria that will be asked to each participant group.

Table 2. Round 1 rating domains by participant group for each indicator.

Participant group and rating assessment criteria	Questions	Rating scale
Patients and caregivers		
Importance to patients and caregivers (face validity)	How important is this indicator to patients and caregivers affected by MLTCs^a?	Nine-point Likert scale for importance, where 1 indicates “extremely unimportant” and 9 indicates “extremely important”
Comprehensiveness or clarity	Is the definition of this indicator clear enough to understand?	Yes, no, or free text
Redundancy	To what extent does this indicator reflect current routine clinical practice?	Nine-point Likert scale, where 1 indicates “does not reflect routine practice,” 5 indicates “partially reflects routine practice,” and 9 indicates “very accurately reflects routine practice”
Health care professionals		
Importance	How important is this indicator to patients and caregivers affected by MLTC?	Nine-point Likert scale for importance
Content validity	How important is this indicator to measure high quality of care related to its targeted domain of care (eg, respectful care)?	Nine-point Likert scale for importance
Face validity	How important is this indicator to primary care clinical activities or practice?	Nine-point Likert scale for importance
Redundancy	- To what extent does this indicator reflect current routine clinical practice? - Are there existing indicators measuring the same domain of care quality currently in use in your practice?	- Yes, no, or free text - Nine-point Likert scale for routine practice
Comprehensiveness or clarity	Is the definition of this indicator clear enough to understand?	Yes, no, or free text

^aMLTC: multiple long-term condition.

To minimize redundancy and avoid future indicator burden [42,52], HCPs will also be asked on whether an indicator is already covered by existing metrics, indicators, or digital systems in current practice. To complement this question, we will inquire whether the indicator quality domain is already being implemented in local clinical practice.

Both patients and HCPs will be provided with a question to comment on the clarity of indicator definition. This will be supplemented by the opportunity to provide written comments and to suggest additional indicators based on their expertise [54]. Clarity will be given as a quality check and not a consensus construct; therefore, it will be voted on in a dichotomous manner (yes, no, or needs changing) instead of a Likert-type scale.

Round 1 will be conducted remotely using an online survey platform Qualtrics, following recommendations for web-based RAM panels [33]. Before rating the indicators, participants will receive information on the previous scoping review, qualitative interview findings, and specific instructions on how to rate the indicators in the first pages of the Qualtrics survey [33]. This ensures that all panel members are familiar with the rating and ranking process, as well as common definitions used throughout the survey. At this stage, panelists will be specifically instructed not to consider feasibility or implementation barriers, such as workload constraints, cost implications, or existing local service configurations, when making their ratings.

All participants will be given up to 5 weeks to respond, with automated reminders sent after 2 weeks of inactivity.

Between Rounds 1 and 2: Data Analysis

Consensus thresholds will be informed by RAM, and ratings will be summarized using medians and IQRs for each indicator and each question. Medians and IQRs will be calculated separately for each stakeholder group given the heterogeneous nature of the panel combining lived-experience and professional perspectives.

Disagreement will be defined as evidence of polarization, whereby at least one-third of the total panel rates an indicator in the 1-3 region and at least one-third rates the same indicator in the 7-9 region [33].

The IQR will be used as a measure of dispersion to assess the degree of agreement among panel members and will not be used as a strict threshold for consensus [34]. Lower IQR values (eg, IQR ≤2) suggest greater consensus, and high IQR values (eg, IQR ≥4) indicate wider dispersion. In the absence of disagreement, final indicator decisions will be based on median scores as follows [33,34]:

- Indicator will be *accepted* and proceed straight to round 3: if both stakeholder-group median ratings are between 7 and 9 across rating assessment criteria questions, with no disagreement.
- Indicator will be classified as “*uncertain*” and *retained for round 2 discussion*: if any stakeholder-group median is 4 to 6, if disagreement is present, or if stakeholder groups are discordant, for example, if one group rates the indicator on a median between 7 to 9 and the other group rates it below a median of 6 in any question.

- Indicator will be *rejected and excluded* from subsequent rounds: if both stakeholder-group medians rating are between 1 and 3, with no disagreement.

Written comments from free-text questions will be analyzed through content analysis methods to identify refinement needed in the definitions provided or overlap with existing systems indicating redundancy [45]. In line with recommendations for assessing the quality of Delphi studies, the volume and depth of qualitative responses (eg, presence of substantive, explanatory comments) will be examined as an indicator of participant engagement [45,55]. Based on these comments, indicators may be merged, reworded, or excluded. Indicators will be retained for round 2 discussion if free-text comments indicate the need for discussion, conceptual overlap between indicators, or limited added value to lived experiences or clinical practice, even if they are rated as highly important by any stakeholder group. Sensitivity analyses will be conducted to assess the robustness of the findings, including comparisons of indicator ratings between patients and HCPs. Where relevant, to assess potential nonresponse bias, a post hoc Mann-Whitney *U* test comparing early and late responders, using median indicator ratings, will be conducted [56].

To address context-specific operationalization of indicators, the research team will develop technical specifications for indicators that are accepted or marked as “uncertain.” This will include the development of provisional numerators, denominators, exclusions, and data sources from which indicators can be measured in current primary care practices in England (Table S3 in [Multimedia Appendix 1](#)). For indicators proceeding to round 2 discussion, these technical specifications will be used to support the consensus meeting discussion.

Round 2: Consensus Meeting

Controlled feedback, a classic feature of Delphi-type studies, will be provided to each panel member [44]. An individual feedback report will be sent 1 to 2 weeks prior to the round 2 consensus meeting. This will include participants’ round 1 ratings, anonymized comments from participants, and summary statistics (median and IQR) for each indicator by stakeholder group, presented using a color-coded system to indicate agreement status (Table S2 in [Multimedia Appendix 1](#)).

For patient participants, a lay summary will be coproduced with our community partners including participants’ own responses, the overall panel median without IQRs, and relevant free-text comments. Individual prior scores will not be disclosed to the wider group maintaining data anonymously.

The round 2 consensus meeting is a characteristic of RAM [33] and will consist of a fully online or hybrid consensus workshop, across 1 or 2 days, depending on participant availability [57]. Given the large initial set of candidate

indicators and to reduce participant attrition, this meeting will review only those indicators marked as “uncertain” or demonstrating conceptual ambiguity or redundancy through free-text comments made in round 1.

At the start of the consensus meeting, all indicators marked as uncertain will be presented sequentially to the expert panel, each including a definition, proposed numerator, denominator, and data sources (Table S3 in [Multimedia Appendix 1](#)). Participants will independently rate each indicator in real time using an audience response system (eg, Teams voting polls), selecting one of three options: (1) accept indicator as an important measure of high quality of care for those with MLTCs, (2) indicator requires discussion before accepting or rejecting, (3) discard indicator as a not important measure of high quality of care. No discussion will take place during this silent voting process, ensuring balanced participation and preventing dominance of individual viewpoints, particularly between patients and health care providers.

Once all the panel votes on these indicators, predefined thresholds will be used to prioritize indicators discussion process within the meeting. Indicators receiving $\geq 60\%$ agreement for accept will not be discussed further at this stage and will proceed directly to the final rerating round 3. Indicators receiving $\geq 40\%$ “requires discussion” votes will be prioritized for structured discussion. Indicators receiving $>20\%$ “do not retain” votes will be discussed explicitly before any decision is made about removal. If indicators receive $\geq 60\%$ reject votes, then the indicator is rejected without further discussion.

The panel will be asked to comment on clarity of definitions, wording refinements, and conceptual overlap with other indicators. HCP participants will be asked to consider the feasibility of operationalization, including the proposed numerator, denominator, and data sources. Patient and caregiver participants will be asked whether the proposed measure, including the numerator, reflects their lived experience and captures what matters to them in care.

Discussions will be audio-recorded and transcribed verbatim for analysis. All participants will have provided written informed consent before the meeting. Transcripts will be analyzed using content analysis, and all proposed modifications, decisions, and rationales will be documented to ensure transparency and reproducibility. Final consensus and prioritization will be determined in round 3.

Round 3 Final Rerating and Consensus

Before round 3, participants will be sent via e-mail a feedback report summarizing the main changes made to the indicators, including the rationale for these decisions based on the meeting’s discussions.

During round 3, both stakeholder groups will rerate any indicators that were merged or revised as part of round 2

discussions. This will be conducted using the same web-based platform and 9-point Likert scale and consensus criteria as in round 1. Panelists will be asked to reconsider their ratings based on group discussion and feedback.

To address the operationalization of the indicator set, round 3 will also include specific questions on measurability and feasibility of all indicators reaching round 3 (Table 3). Patient participants will not be asked to rate these measurability and feasibility questions as advised by our CPs.

Table 3. Round 3 additional feasibility rating domains for expert health care professional participants.

Participant group and rating assessment criteria	Questions	Rating scale
Health care participants		
Interpretability and Measurability	- Do you agree that the proposed numerator, denominator, exclusions, and data source are appropriate for measuring this indicator? - Free-text comments.	- Nine-point Likert scale, where 1 indicates “Strongly disagree,” 5 indicates “Partly agree,” and 9 indicates “Strongly agree.” - Please suggest any changes to the numerator, denominator, exclusions, or data source.
Feasibility	- How feasible would it be to measure this indicator using data currently available in routine primary care records or administrative systems? - Free-text comments.	- Nine-point Likert scale, where 1 indicates “Not feasible” (would require substantial EHR^a changes), 5 indicates “Partly feasible” (some data available but some changes are needed), and 9 indicates “Highly feasible” (can be measured using routinely recorded data with minimal additional burden). - Please describe any data source, coding, EHR^a template, workflow, or implementation issue that may affect measurement.

^aEHR: electronic health record.

Feasibility is being evaluated at this stage and separately from importance questions to distinguish indicators that are conceptually important from those that are likely to be immediately measurable or implementable in routine practice. Following round 3 results, the final set of indicators will be categorized according to their likely measurement readiness, following a red/amber/green system: those potentially measurable using existing routine primary care data (green), those requiring some changes to existing local templates and electronic health records (amber), and those requiring extensive changes and classed as difficult to change existing templates (red). This classification will guide subsequent feasibility testing and help assess the transferability of the indicator set across different English primary care settings.

At the conclusion of the RAM study, panel members will be invited to provide feedback on the process and the acceptability of the final indicator set [53].

Ethical Considerations

This study was conducted in accordance with the principles of the Declaration of Helsinki. Ethical approval was obtained through the Imperial College Research Ethics Committee, acting as the institutional ethics review board, and administered by the Imperial College Research Governance and Integrity Team, with Head of Department approval on April 24, 2025 (Imperial College Research Ethics Committee reference 7594295). All participants will receive an information sheet and will be asked to sign a consent form. Survey data from Qualtrics will be collected in an anonymized format as per the platform’s inbuilt confidentiality settings. The consensus meeting recordings will be stored on secure,

password-protected institutional servers and accessible only to authorized members of this research team. Recordings from consensus meetings will be transcribed, pseudonymized, checked for accuracy, and deleted after transcription and data analysis. Participants will be able to withdraw at any time, including between rounds. After their data are anonymized and synthesized into group-level findings, individual data may no longer be identifiable for removal. Between rounds, during participant feedback reports, only anonymized group-level findings will be reported to protect participant confidentiality.

Results

This study received funding in June 2024. The scoping review component was finalized in January 2026 and published in May 2026. Qualitative interview data collection took place between May 2025 and July 2025, with data analysis concluded in December 2025. The associated manuscript reporting the qualitative findings is currently under review elsewhere.

As of July 2026, the RAM consensus process has a core group of 4 community advisory members, who will participate in all stages of this study. Additionally, 10 patient and caregiver participants and 12 HCPs have been consented. Data analysis and results from all 3 planned RAND/UCLA rounds are expected to be published in February 2027.

Discussion

Anticipated Principal Findings

This protocol describes a planned co-designed approach to develop primary care quality indicators for people living with MLTCs. This study will generate quality indicators that are grounded in published evidence and informed by patient and caregiver priorities. Through a 3-round RAM consensus process, indicators will be rated by both patients, caregivers, and expert health care professionals on importance, existing competing indicators, and feasibility on proposed numerators, denominators, and contemporary data sources embedded in existing primary care electronic health records. The final set of indicators will require further testing and validation in routine GP primary care settings in England.

Comparison With Prior Work

Previous work has established important methodological foundations for developing quality indicators in primary care; however, their applicability to MLTC remains limited as they commonly focus on disease-specific care [58], older populations [26,59], or prescribing outputs [16].

A second limitation of previous indicator development is the variable and often limited involvement of patients and caregivers. A systematic review by Kötter et al [60] on the involvement of patients in quality of care indicator development found scarce involvement of patients in panels largely led by professional and technical professionals. Although patients are increasingly involved in consensus panels, they often remain numerically outnumbered by professional participants, with minimal reporting on ethnicity and socioeconomic characteristics [61]. Even when patients are actively involved, patient-prioritized aspects of care may be dismissed because they are perceived as too difficult to measure. For example, Pohontsch et al [62] found that aspects of care considered important to patients were often categorized as too “generic” in contrast to guideline-driven “hard outcome” metrics. Indeed, disease-specific indicators and clinical outcome measures remain important for monitoring population health and evidence-based care. However, especially in MLTCs, these measures need to coexist with indicators that capture microlevel interactions between patients and professionals, meso-level organizational arrangements such as access and coordination, and macro-level system conditions such as workforce capacity, digital infrastructure, and social support [63]. Over-reliance on incentivized disease-specific metrics risks reinforcing checklist-driven care and may contribute to changes in relational continuity and trust between HCPs and patients [64].

A further challenge concerns transferability across clinical, organizational, and national contexts. Although several quality indicators have been developed in the United States [22], Canada [23], and Germany [65], fewer are directly developed specifically for the UK care setting [16]. A systematic review of European studies on quality indicators for hospital readmission rates highlighted that indicators

developed in one clinical, organizational, or national context cannot be assumed to transfer directly to another without attention to local pathways, data infrastructure, professional workflows, and patient priorities [66]. In the United Kingdom, the NICE multimorbidity quality standard [17] remains a useful starting point, but it is brief, with only four priority statements to function as a framework for assessing quality of care for people with MLTCs. When compared with broader patient-centered care quality indicator frameworks, its limitations are evident. The work of Santana et al [28] on person-centered care quality indicators identified a much wider set of 26 indicators across structure, process, outcome, and global assessment categories. The MULTiqual study by Schulze et al [59] further exposes this narrowness by validating a core set of 12 GP-reported and 7 patient-reported indicators in Germany’s primary care system, recognizing that quality in MLTCs must be assessed through professional report, patient experience, and documented clinical processes. Although not aiming to develop quality indicators, the European SELFIE conceptual framework for integrated care for multimorbidity also exposes NICE’s limitations. SELFIE, led by Leijten et al [67], organized care around the 6 adapted World Health Organization health-system components: service delivery, leadership and governance, workforce, financing, technologies and medical products, and information and research, across micro, meso, and macro levels. Therefore, where NICE offers a concise set of minimum standards, these frameworks suggest that quality in MLTCs should move beyond a small number of care-process statements and toward a layered model that includes patient-reported experience, professional assessment, organizational capability, information infrastructure, and system-level integration.

Strengths and Limitations

One of this study’s strengths is the central role of patients and caregivers in indicator development. Rather than being invited only to comment on final outputs, patients and carers contribute to the conceptual framing of quality and to the design of the consensus process. Given the well-documented differences in priorities between patients and HCPs [13,68], particularly in MLTC care [69,70], the inclusion of patients and caregivers in consensus processes is essential to ensure that indicators reflect priorities that are meaningful to those receiving care, as well as those delivering it [41]. The study is also designed with future implementation in mind, including feasibility assessment of embedding proposed indicators in existing pathways and systems during rounds 2 and 3.

Nevertheless, this study also has limitations. First, there are no well-established quality indicators for MLTCs in primary care that are not specific to individual risk factors or conditions. The heterogeneity of the published literature developing or evaluating quality indicators in MLTC was scarce, with only 1 study developing well-explained methodologies and evaluating the validity of an indicator set [59,65]. Reliance on qualitative work means that substantial interpretive work is required to translate broad themes and subthemes into discrete, measurable candidate indicators. Consequently, several indicators may conceptually sit

within more than one quality domain. Although this will be addressed through iterative discussion with the study team, core participants, and structured consensus methods, some overlap is likely to remain.

Finally, Delphi-type methods are inherently judgment-based, and particularly in MLTCs where consensus on its definition is still up for debate [71], different panels might reach somewhat different conclusions when presented with the same evidence, although some argue that this is not always the case [47]. To strengthen credibility and relevance, this study will include a broad group of knowledge users and apply a transparent rating process to combine empirical evidence with stakeholder expertise.

Future Directions

The final indicator set deriving from this RAM study should be viewed as a consensus-based starting point that requires further validation, feasibility testing, and real-world evaluation rather than immediate implementation. While some indicators are likely to be more readily measurable, particularly those aligned with existing systems and quality statements, others will require new data capture or refinement once tested in practice. For example, indicators measuring medication reviews are already partially in place in many primary care electronic health records in the UK as part of

NICE quality standards and therefore expected to be easier to achieve [17,72].

Future work will test the feasibility, transferability, and implementation of the final indicator set in routine GP settings in England. A Freedom of Information request to Integrated Care Boards in England has been initiated to understand how current quality indicators beyond nationally mandated QOF are currently captured, monitored, and used in practice, including data systems, governance, dashboards, and reporting processes that support local quality monitoring. Pooled findings from this request will identify which existing data sources and infrastructures could be used to operationalize the proposed indicators and where local variation may affect implementation across primary care services in England [73]. A freedom of information request rather than a survey or gray literature search was selected because locally developed dashboards and quality monitoring frameworks are not routinely described in the academic literature or publicly available. Surveys of policymakers and system leaders may be limited by low response rates [74].

Following this, qualitative work with local improvement leaders and electronic health record developers will explore how indicators can be operationalized, automated within routine data systems, and integrated into practice.

Acknowledgments

We would like to thank the patient and public contributors, the Patient Experience Center staff for their engagement with our study, our interview participants, and health care stakeholders who so far already informed the development of the candidate indicator set and planned Delphi materials. ChatGPT was used only for refinement and correction of grammar and spelling errors to improve clarity of language. All scientific content, interpretation, and final revisions were made by the authors.

Funding

This research project was funded by the National Institute for Health and Care Research Imperial Biomedical Research Centre. The grant reference number is NIHR203323. The funder had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript. The views expressed are those of the authors and not necessarily those of the National Institute for Health and Care Research or the Department of Health and Social Care.

Data Availability

The datasets generated and/or analyzed during the current study are not publicly available due to confidentiality but are available from the corresponding author on reasonable request.

Authors' Contributions

Aiding formal analysis: TB, AR, AH, DB, AE, LD

Conceptualization: ST, LD

Data curation and collection: ST

Formal analysis: ST, TB, AR, AH, DB, LD

Funding acquisition: LD

Investigation: ST

Methodology: ST

Project administration: ST

Software (Qualtrics, Excel): ST

Supervision: TB, LD

Validation: ST

Visualization: ST

Writing – original draft: ST

Writing – review and editing: ST, TB, AR, AH, DB, AE, LD

Conflicts of Interest

None declared.

Multimedia Appendix 1

Quality of care indicator development additional information.

[[DOCX File \(Microsoft Word File\), 28 KB-Multimedia Appendix 1](#)]

Checklist 1

GRIPP 2-Short Form.

[[DOCX File \(Microsoft Word File\), 32 KB-Checklist 1](#)]

References

1. Kingston A, Comas-Herrera A, Jagger C, MODEM project. Forecasting the care needs of the older population in England over the next 20 years: estimates from the Population Ageing and Care Simulation (PACSim) modelling study. *Lancet Public Health*. Sep 2018;3(9):e447-e455. [doi: [10.1016/S2468-2667\(18\)30118-X](#)] [Medline: [30174210](#)]
2. Hajat C, Stein E. The global burden of multiple chronic conditions: a narrative review. *Prev Med Rep*. Dec 2018;12:284-293. [doi: [10.1016/j.pmedr.2018.10.008](#)] [Medline: [30406006](#)]
3. Chowdhury SR, Das DC, Sunna TC, Beyene J, Hossain A. Global and regional prevalence of multimorbidity in the adult population in community settings: a systematic review and meta-analysis. *EClinicalMedicine*. Mar 2023;57:101860. [doi: [10.1016/j.eclinm.2023.101860](#)] [Medline: [36864977](#)]
4. France EF, Wyke S, Gunn JM, Mair FS, McLean G, Mercer SW. Multimorbidity in primary care: a systematic review of prospective cohort studies. *Br J Gen Pract*. Apr 2012;62(597):e297-307. [doi: [10.3399/bjgp12X636146](#)] [Medline: [22520918](#)]
5. Makovski TT, Schmitz S, Zeegers MP, Stranges S, van den Akker M. Multimorbidity and quality of life: systematic literature review and meta-analysis. *Ageing Res Rev*. Aug 2019;53:100903. [doi: [10.1016/j.arr.2019.04.005](#)] [Medline: [31048032](#)]
6. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. *The Lancet*. Jul 2012;380(9836):37-43. [doi: [10.1016/S0140-6736\(12\)60240-2](#)]
7. Simpson G, Morrison L, Santer M, Hijryana M, Farmer A, Dambha-Miller H. Perceptions and experiences of living with and providing care for multimorbidity: a qualitative interview study. *J Multimorb Comorb*. 2024;14:26335565241240820. [doi: [10.1177/26335565241240820](#)] [Medline: [38529048](#)]
8. Paddison CAM, Abel GA, Roland MO, Elliott MN, Lyratzopoulos G, Campbell JL. Drivers of overall satisfaction with primary care: evidence from the English General Practice Patient Survey. *Health Expect*. Oct 2015;18(5):1081-1092. [doi: [10.1111/hex.12081](#)] [Medline: [23721257](#)]
9. Campbell SM, Braspenning J, Hutchinson A, Marshall M. Research methods used in developing and applying quality indicators in primary care. *Qual Saf Health Care*. Dec 2002;11(4):358-364. [doi: [10.1136/qhc.11.4.358](#)] [Medline: [12468698](#)]
10. Lee JE, Lee J, Shin R, Oh O, Lee KS. Treatment burden in multimorbidity: an integrative review. *BMC Prim Care*. Sep 28, 2024;25(1):352. [doi: [10.1186/s12875-024-02586-z](#)] [Medline: [39342121](#)]
11. Holland E, Matthews K, Macdonald S, et al. The impact of living with multiple long-term conditions (multimorbidity) on everyday life—a qualitative evidence synthesis. *BMC Public Health*. Dec 18, 2024;24(1):3446. [doi: [10.1186/s12889-024-20763-8](#)] [Medline: [39696210](#)]
12. Tran VT, Riveros C, Péan C, Czarnobroda A, Ravaud P. Patients' perspective on how to improve the care of people with chronic conditions in France: a citizen science study within the ComPaRe e-cohort. *BMJ Qual Saf*. Nov 2019;28(11):875-886. [doi: [10.1136/bmjqs-2018-008593](#)] [Medline: [31015376](#)]
13. Sathanapally H, Sidhu M, Fahami R, et al. Priorities of patients with multimorbidity and of clinicians regarding treatment and health outcomes: a systematic mixed studies review. *BMJ Open*. Feb 12, 2020;10(2):e033445. [doi: [10.1136/bmjopen-2019-033445](#)] [Medline: [32051314](#)]
14. Valderas JM, Gangannagaripalli J, Nolte E, et al. Quality of care assessment for people with multimorbidity. *J Intern Med*. Mar 2019;285(3):289-300. [doi: [10.1111/joim.12881](#)] [Medline: [30719790](#)]
15. Ho JW, Kuluski K, Im J. "It's a fight to get anything you need"—Accessing care in the community from the perspectives of people with multimorbidity. *Health Expect*. Dec 2017;20(6):1311-1319. [doi: [10.1111/hex.12571](#)] [Medline: [28544493](#)]

16. Chambers D, Mawson R, Mettle-Nunoo J, Sutton A, Booth A. A systematic review of international performance indicators and metrics relevant to UK general practice. *BMJ Open Qual.* Oct 15, 2025;14(4):e003477. [doi: [10.1136/bmjopen-2025-003477](https://doi.org/10.1136/bmjopen-2025-003477)] [Medline: [41093614](https://pubmed.ncbi.nlm.nih.gov/41093614/)]
17. CKS via NICE is only available to eligible users within the UK. National Institute for Health and Care Excellence (NICE). 2023. URL: <https://cks.nice.org.uk/topics/multimorbidity/goals-outcome-measures/nice-quality-standards/> [Accessed 2023-07-29]
18. Quality and outcomes framework guidance for 2026/27. NHS England; 2026. URL: <https://www.england.nhs.uk/wp-content/uploads/2026/03/prn02356-quality-outcomes-framework-guidance-2026-27-july-update.pdf> [Accessed 2026-07-29]
19. Ferris R, Blaum C, Kiwak E, et al. Perspectives of patients, clinicians, and health system leaders on changes needed to improve the health care and outcomes of older adults with multiple chronic conditions. *J Aging Health.* Jun 2018;30(5):778-799. [doi: [10.1177/0898264317691166](https://doi.org/10.1177/0898264317691166)] [Medline: [28553806](https://pubmed.ncbi.nlm.nih.gov/28553806/)]
20. Morris RL, Sanders C, Kennedy AP, Rogers A. Shifting priorities in multimorbidity: a longitudinal qualitative study of patient's prioritization of multiple conditions. *Chronic Illn.* Jun 2011;7(2):147-161. [doi: [10.1177/1742395310393365](https://doi.org/10.1177/1742395310393365)] [Medline: [21343220](https://pubmed.ncbi.nlm.nih.gov/21343220/)]
21. Hanefeld J, Powell-Jackson T, Balabanova D. Understanding and measuring quality of care: dealing with complexity. *Bull World Health Organ.* May 1, 2017;95(5):368-374. [doi: [10.2471/BLT.16.179309](https://doi.org/10.2471/BLT.16.179309)] [Medline: [28479638](https://pubmed.ncbi.nlm.nih.gov/28479638/)]
22. Bayliss EA, McQuillan DB, Ellis JL, et al. Using electronic health record data to measure care quality for individuals with multiple chronic medical conditions. *J Am Geriatr Soc.* Sep 2016;64(9):1839-1844. [doi: [10.1111/jgs.14248](https://doi.org/10.1111/jgs.14248)] [Medline: [27385077](https://pubmed.ncbi.nlm.nih.gov/27385077/)]
23. Manalili K, Scott CM, O'Beirne M, Hemmelgarn BR, Santana MJ. Informing the implementation and use of person-centred quality indicators: a mixed methods study on the readiness, barriers and facilitators to implementation in Canada. *BMJ Open.* Aug 25, 2022;12(8):e060441. [doi: [10.1136/bmjopen-2021-060441](https://doi.org/10.1136/bmjopen-2021-060441)] [Medline: [36008077](https://pubmed.ncbi.nlm.nih.gov/36008077/)]
24. Doktorchik C, Manalili K, Jolley R, et al. Identifying Canadian patient-centred care measurement practices and quality indicators: a survey. *CMAJ Open.* 2018;6(4):E643-E650. [doi: [10.9778/cmajo.20170143](https://doi.org/10.9778/cmajo.20170143)] [Medline: [30563920](https://pubmed.ncbi.nlm.nih.gov/30563920/)]
25. Derendorf L, Stock S, Simic D, Lemmen C. Developing quality indicators for cross-sectoral psycho-oncology in Germany: combining the RAND/UCLA appropriateness method with a Delphi technique. *BMC Health Serv Res.* Jun 8, 2023;23(1):599. [doi: [10.1186/s12913-023-09604-3](https://doi.org/10.1186/s12913-023-09604-3)] [Medline: [37291536](https://pubmed.ncbi.nlm.nih.gov/37291536/)]
26. Correia RH, Siu HYH, Vanstone M, Jones A, Gopaul A, Costa AP. Development of practice-based quality indicators for the primary care of older adults: a RAND/UCLA Appropriateness Method study protocol. *BMJ Open.* Sep 12, 2023;13(9):e072232. [doi: [10.1136/bmjopen-2023-072232](https://doi.org/10.1136/bmjopen-2023-072232)] [Medline: [37699633](https://pubmed.ncbi.nlm.nih.gov/37699633/)]
27. Hartveit M, Vanhaecht K, Thorsen O, Biringer E, Haug K, Aslaksen A. Quality indicators for the referral process from primary to specialised mental health care: an explorative study in accordance with the RAND appropriateness method. *BMC Health Serv Res.* Jan 3, 2017;17(1):4. [doi: [10.1186/s12913-016-1941-1](https://doi.org/10.1186/s12913-016-1941-1)] [Medline: [28049470](https://pubmed.ncbi.nlm.nih.gov/28049470/)]
28. Santana MJ, Manalili K, Zelinsky S, et al. Improving the quality of person-centred healthcare from the patient perspective: development of person-centred quality indicators. *BMJ Open.* Oct 29, 2020;10(10):e037323. [doi: [10.1136/bmjopen-2020-037323](https://doi.org/10.1136/bmjopen-2020-037323)] [Medline: [33122312](https://pubmed.ncbi.nlm.nih.gov/33122312/)]
29. Saust LT, Siersma VD, Bjerrum L, Hansen MP. Development of quality indicators for the diagnosis and treatment of urinary tract infections in general practice: a RAND appropriateness method. *BMJ Open Qual.* May 2023;12(2):e002156. [doi: [10.1136/bmjopen-2022-002156](https://doi.org/10.1136/bmjopen-2022-002156)] [Medline: [37230738](https://pubmed.ncbi.nlm.nih.gov/37230738/)]
30. Smets M, Smeets M, Van den Bulck S, Janssens S, Aertgeerts B, Vaes B. Defining quality indicators for heart failure in general practice. *Acta Cardiol.* Aug 2019;74(4):291-298. [doi: [10.1080/00015385.2018.1492659](https://doi.org/10.1080/00015385.2018.1492659)] [Medline: [30261830](https://pubmed.ncbi.nlm.nih.gov/30261830/)]
31. Van den Bulck SA, Vankrunkelsven P, Goderis G, et al. Developing quality indicators for chronic kidney disease in primary care, extractable from the electronic medical record. A RAND-modified Delphi method. *BMC Nephrol.* May 5, 2020;21(1):161. [doi: [10.1186/s12882-020-01788-8](https://doi.org/10.1186/s12882-020-01788-8)] [Medline: [32370742](https://pubmed.ncbi.nlm.nih.gov/32370742/)]
32. Dudley L, Mamdoo P, Naidoo S, Muzigaba M. Towards a harmonised framework for developing quality of care indicators for global health: a scoping review of existing conceptual and methodological practices. *BMJ Health Care Inform.* Jan 2022;29(1):e100469. [doi: [10.1136/bmjhci-2021-100469](https://doi.org/10.1136/bmjhci-2021-100469)] [Medline: [35078776](https://pubmed.ncbi.nlm.nih.gov/35078776/)]
33. Fitch KB, Bernstein SJ, Aguilar MD, et al. The RAND/UCLA appropriateness method user's manual. RAND Corporation; 2001. URL: https://www.rand.org/content/dam/rand/pubs/monograph_reports/2011/MR1269.pdf [Accessed 2026-07-29]
34. Niederberger M, Schifano J, Deckert S, et al. Delphi studies in social and health sciences—recommendations for an interdisciplinary standardized reporting (DELPHISTAR). Results of a Delphi study. *PLOS One.* 2024;19(8):e0304651. [doi: [10.1371/journal.pone.0304651](https://doi.org/10.1371/journal.pone.0304651)] [Medline: [39186713](https://pubmed.ncbi.nlm.nih.gov/39186713/)]

35. Rodríguez-Labajos L, Thomson CS, O'Brien G. Applying constructivist grounded theory in co-production research: a case study exploring its potential and lessons for construction management research. *Constr Manag Econ*. May 4, 2021;39(5):369-382. [doi: [10.1080/01446193.2021.1894654](https://doi.org/10.1080/01446193.2021.1894654)]
36. Kaushik V, Walsh CA. Pragmatism as a research paradigm and its implications for social work research. *Soc Sci*. 2019;8(9):255. [doi: [10.3390/socsci8090255](https://doi.org/10.3390/socsci8090255)]
37. Ormerod RJ. Pragmatism in professional practice. *Syst Res Behav Sci*. Nov 2021;38(6):797-816. [doi: [10.1002/sres.2739](https://doi.org/10.1002/sres.2739)]
38. Ryan G. Introduction to positivism, interpretivism and critical theory. *Nurse Res*. Mar 16, 2018;25(4):14-20. [doi: [10.7748/nr.2018.e1466](https://doi.org/10.7748/nr.2018.e1466)] [Medline: [29546962](https://pubmed.ncbi.nlm.nih.gov/29546962/)]
39. Improving how we partner with patients, carers and the public. NIHR (National Institute for Health and Care Research). 2024. URL: <https://www.nihr.ac.uk/about-us/what-we-do/improving-partner-patients-carers-public> [Accessed 2026-07-29]
40. Staniszewska S, Brett J, Simera I, et al. GRIPP2 reporting checklists: tools to improve reporting of patient and public involvement in research. *BMJ*. Aug 2, 2017;358:j3453. [doi: [10.1136/bmj.j3453](https://doi.org/10.1136/bmj.j3453)] [Medline: [28768629](https://pubmed.ncbi.nlm.nih.gov/28768629/)]
41. Santana MJ, Ahmed S, Lorenzetti D, et al. Measuring patient-centred system performance: a scoping review of patient-centred care quality indicators. *BMJ Open*. Jan 7, 2019;9(1):e023596. [doi: [10.1136/bmjopen-2018-023596](https://doi.org/10.1136/bmjopen-2018-023596)] [Medline: [30617101](https://pubmed.ncbi.nlm.nih.gov/30617101/)]
42. Schang L, Blotenberg I, Boywitt D. What makes a good quality indicator set? A systematic review of criteria. *Int J Qual Health Care*. Jul 31, 2021;33(3):mzab107. [doi: [10.1093/intqhc/mzab107](https://doi.org/10.1093/intqhc/mzab107)] [Medline: [34282841](https://pubmed.ncbi.nlm.nih.gov/34282841/)]
43. Tavares S, Beaney T, Reisberg A, et al. The content and processes of patient-derived quality of care indicators for people living with multiple long-term conditions (MLTC): a scoping review. *J Multimorb Comorb*. 2026;16:26335565261451686. [doi: [10.1177/26335565261451686](https://doi.org/10.1177/26335565261451686)] [Medline: [42153096](https://pubmed.ncbi.nlm.nih.gov/42153096/)]
44. Nasa P, Jain R, Juneja D. Delphi methodology in healthcare research: how to decide its appropriateness. *World J Methodol*. Jul 20, 2021;11(4):116-129. [doi: [10.5662/wjmv11.i4.116](https://doi.org/10.5662/wjmv11.i4.116)] [Medline: [34322364](https://pubmed.ncbi.nlm.nih.gov/34322364/)]
45. Beiderbeck D, Frevel N, von der Gracht HA, Schmidt SL, Schweitzer VM. Preparing, conducting, and analyzing Delphi surveys: cross-disciplinary practices, new directions, and advancements. *MethodsX*. 2021;8:101401. [doi: [10.1016/j.mex.2021.101401](https://doi.org/10.1016/j.mex.2021.101401)] [Medline: [34430297](https://pubmed.ncbi.nlm.nih.gov/34430297/)]
46. Trotter RT. Qualitative research sample design and sample size: resolving and unresolved issues and inferential imperatives. *Prev Med*. Nov 2012;55(5):398-400. [doi: [10.1016/j.ypmed.2012.07.003](https://doi.org/10.1016/j.ypmed.2012.07.003)] [Medline: [22800684](https://pubmed.ncbi.nlm.nih.gov/22800684/)]
47. Boel A, Navarro-Compán V, Landewé R, van der Heijde D. Two different invitation approaches for consecutive rounds of a Delphi survey led to comparable final outcome. *J Clin Epidemiol*. Jan 2021;129:31-39. [doi: [10.1016/j.jclinepi.2020.09.034](https://doi.org/10.1016/j.jclinepi.2020.09.034)] [Medline: [32991995](https://pubmed.ncbi.nlm.nih.gov/32991995/)]
48. Donabedian A. The quality of care. How can it be assessed? *JAMA*. 1988;260(12):1743-1748. [doi: [10.1001/jama.260.12.1743](https://doi.org/10.1001/jama.260.12.1743)] [Medline: [3045356](https://pubmed.ncbi.nlm.nih.gov/3045356/)]
49. Terwee CB, Prinsen CAC, Chiarotto A, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Qual Life Res*. May 2018;27(5):1159-1170. [doi: [10.1007/s11136-018-1829-0](https://doi.org/10.1007/s11136-018-1829-0)] [Medline: [29550964](https://pubmed.ncbi.nlm.nih.gov/29550964/)]
50. Quentin W, Partanen VM, Brownwood I, Klazinga N. Measuring healthcare quality. In: Busse R, Klazinga N, Panteli D, Quentin W, editors. *Improving Healthcare Quality in Europe: Characteristics, Effectiveness and Implementation of Different Strategies*. European Observatory on Health Systems and Policies; 2019:39-63. URL: <https://www.ncbi.nlm.nih.gov/books/NBK549260/> [Accessed 2026-07-29]
51. Busse R, Klazinga N, Panteli D, Quentin W, editors. *Improving Healthcare Quality in Europe: Characteristics, Effectiveness and Implementation of Different Strategies*. European Observatory on Health Systems and Policies; 2019. ISBN: 9789289051750
52. Building a roadmap from patient-reported outcome measures to patient-reported outcome performance measures. National Quality Forum (NQF); 2021. URL: <https://digitalassets.jointcommission.org/api/public/content/c0c7bed4c69b41eb9c782bf1fbc8ed3e?v=09abe566> [Accessed 2026-07-29]
53. Sparks JB, Klamerus ML, Caverly TJ, et al. Planning and reporting effective web-based RAND/UCLA appropriateness method panels: literature review and preliminary recommendations. *J Med Internet Res*. Aug 26, 2022;24(8):e33898. [doi: [10.2196/33898](https://doi.org/10.2196/33898)] [Medline: [36018626](https://pubmed.ncbi.nlm.nih.gov/36018626/)]
54. Banno M, Tsujimoto Y, Kataoka Y. Reporting quality of the Delphi technique in reporting guidelines: a protocol for a systematic analysis of the EQUATOR Network Library. *BMJ Open*. Apr 3, 2019;9(4):e024942. [doi: [10.1136/bmjopen-2018-024942](https://doi.org/10.1136/bmjopen-2018-024942)] [Medline: [30948575](https://pubmed.ncbi.nlm.nih.gov/30948575/)]
55. Roßmann B, Canzaniello A, von der Gracht H, Hartmann E. The future and social impact of Big Data Analytics in Supply Chain Management: results from a Delphi study. *Technol Forecast Soc Change*. May 2018;130:135-149. [doi: [10.1016/j.techfore.2017.10.005](https://doi.org/10.1016/j.techfore.2017.10.005)]

56. Sheikh K, Mattingly S. Investigating non-response bias in mail surveys. *J Epidemiol Community Health*. Dec 1981;35(4):293-296. [doi: [10.1136/jech.35.4.293](https://doi.org/10.1136/jech.35.4.293)] [Medline: [6461711](https://pubmed.ncbi.nlm.nih.gov/6461711/)]
57. Woodyatt CR, Finneran CA, Stephenson R. In-person versus online focus group discussions: a comparative analysis of data quality. *Qual Health Res*. May 2016;26(6):741-749. [doi: [10.1177/1049732316631510](https://doi.org/10.1177/1049732316631510)] [Medline: [26935719](https://pubmed.ncbi.nlm.nih.gov/26935719/)]
58. Schneberk T, Bolshakova M, Sloan K, et al. Quality indicators for high-need patients: a systematic review. *J Gen Intern Med*. Sep 2022;37(12):3147-3161. [doi: [10.1007/s11606-022-07454-z](https://doi.org/10.1007/s11606-022-07454-z)] [Medline: [35260956](https://pubmed.ncbi.nlm.nih.gov/35260956/)]
59. Schulze J, Glassen K, Pohontsch NJ, et al. Measuring the quality of care for older adults with multimorbidity: results of the MULTiqual project. *Gerontologist*. Sep 7, 2022;62(8):1135-1146. [doi: [10.1093/geront/gnac013](https://doi.org/10.1093/geront/gnac013)] [Medline: [35090014](https://pubmed.ncbi.nlm.nih.gov/35090014/)]
60. Kötter T, Schaefer FA, Scherer M, Blozik E. Involving patients in quality indicator development—a systematic review. *Patient Prefer Adherence*. 2013;7:259-268. [doi: [10.2147/PPA.S39803](https://doi.org/10.2147/PPA.S39803)] [Medline: [23569365](https://pubmed.ncbi.nlm.nih.gov/23569365/)]
61. Barrington H, Young B, Williamson PR. Patient participation in Delphi surveys to develop core outcome sets: systematic review. *BMJ Open*. Sep 2, 2021;11(9):e051066. [doi: [10.1136/bmjopen-2021-051066](https://doi.org/10.1136/bmjopen-2021-051066)] [Medline: [34475183](https://pubmed.ncbi.nlm.nih.gov/34475183/)]
62. Pohontsch NJ, Herzberg H, Joos S, Welti F, Scherer M, Blozik E. The professional perspective on patient involvement in the development of quality indicators: a qualitative analysis using the example of chronic heart failure in the German health care setting. *Patient Prefer Adherence*. 2015;9:151-159. [doi: [10.2147/PPA.S74064](https://doi.org/10.2147/PPA.S74064)] [Medline: [25653507](https://pubmed.ncbi.nlm.nih.gov/25653507/)]
63. Valderas JM, Porter I, Martin-Delgado J, et al. Development of the Patient-Reported Indicator Surveys (PaRIS) conceptual framework to monitor and improve the performance of primary care for people living with chronic conditions. *BMJ Qual Saf*. Jul 18, 2025;34(8):529-536. [doi: [10.1136/bmjqs-2024-017301](https://doi.org/10.1136/bmjqs-2024-017301)] [Medline: [39174334](https://pubmed.ncbi.nlm.nih.gov/39174334/)]
64. Lester HE, Hannon KL, Campbell SM. Identifying unintended consequences of quality indicators: a qualitative study. *BMJ Qual Saf*. Dec 2011;20(12):1057-1061. [doi: [10.1136/bmjqs.2010.048371](https://doi.org/10.1136/bmjqs.2010.048371)] [Medline: [21693464](https://pubmed.ncbi.nlm.nih.gov/21693464/)]
65. Schäfer I, Schulze J, Glassen K, et al. Validation of patient- and GP-reported core sets of quality indicators for older adults with multimorbidity in primary care: results of the cross-sectional observational MULTiqual validation study. *BMC Med*. Apr 17, 2023;21(1):148. [doi: [10.1186/s12916-023-02856-0](https://doi.org/10.1186/s12916-023-02856-0)] [Medline: [37069536](https://pubmed.ncbi.nlm.nih.gov/37069536/)]
66. Fischer C, Anema HA, Klazinga NS. The validity of indicators for assessing quality of care: a review of the European literature on hospital readmission rate. *Eur J Public Health*. Aug 2012;22(4):484-491. [doi: [10.1093/eurpub/ckr165](https://doi.org/10.1093/eurpub/ckr165)] [Medline: [22140251](https://pubmed.ncbi.nlm.nih.gov/22140251/)]
67. Leijten FRM, Struckmann V, van Ginneken E, et al. The SELFIE framework for integrated care for multi-morbidity: development and description. *Health Policy*. Jan 2018;122(1):12-22. [doi: [10.1016/j.healthpol.2017.06.002](https://doi.org/10.1016/j.healthpol.2017.06.002)] [Medline: [28668222](https://pubmed.ncbi.nlm.nih.gov/28668222/)]
68. Shah AM, Yan X, Tariq S, Ali M. What patients like or dislike in physicians: analyzing drivers of patient satisfaction and dissatisfaction using a digital topic modeling approach. *Inf Process Manag*. May 2021;58(3):102516. [doi: [10.1016/j.ipm.2021.102516](https://doi.org/10.1016/j.ipm.2021.102516)]
69. Tran VT, Diard E, Ravaud P. Priorities to improve the care for chronic conditions and multimorbidity: a survey of patients and stakeholders nested within the ComPaRe e-cohort. *BMJ Qual Saf*. Jul 2021;30(7):577-587. [doi: [10.1136/bmjqs-2020-011219](https://doi.org/10.1136/bmjqs-2020-011219)] [Medline: [32839207](https://pubmed.ncbi.nlm.nih.gov/32839207/)]
70. Rutten-van Mölken M, Karimi M, Leijten F, et al. Comparing patients' and other stakeholders' preferences for outcomes of integrated care for multimorbidity: a discrete choice experiment in eight European countries. *BMJ Open*. Oct 10, 2020;10(10):e037547. [doi: [10.1136/bmjopen-2020-037547](https://doi.org/10.1136/bmjopen-2020-037547)] [Medline: [33039997](https://pubmed.ncbi.nlm.nih.gov/33039997/)]
71. Beaney T. Is consensus attainable on the definition of multiple long term conditions? *BMJ*. Mar 7, 2024;384:q230. [doi: [10.1136/bmj.q230](https://doi.org/10.1136/bmj.q230)] [Medline: [38453185](https://pubmed.ncbi.nlm.nih.gov/38453185/)]
72. Structured medication reviews and medicines optimisation. NHS England. 2026. URL: <https://www.england.nhs.uk/primary-care/pharmacy/smr/> [Accessed 2026-07-29]
73. Nagel DA, Charlton P, Azar R, Koenig L, Burns K, Kean T. RADAR-ES: a methodological framework for conducting environmental scans in health services delivery research. *J Prim Care Community Health*. 2025;16:21501319251363783. [doi: [10.1177/21501319251363783](https://doi.org/10.1177/21501319251363783)] [Medline: [40819256](https://pubmed.ncbi.nlm.nih.gov/40819256/)]
74. Yale SE, Brown MEL, Byrne MHV. Using freedom of information requests to access novel data sources in health professions education research. *Postgrad Med J*. Apr 22, 2025;101(1195):481-486. [doi: [10.1093/postmj/qgae166](https://doi.org/10.1093/postmj/qgae166)] [Medline: [39591525](https://pubmed.ncbi.nlm.nih.gov/39591525/)]

Abbreviations

CP: community partner

GP: general practice

GRIPP: Guidance for Reporting Involvement of Patients and the Public

HCP: health care professional

MLTC: multiple long-term condition

NICE: National Institute for Health and Care Excellence

QOF: Quality Outcomes Framework

RAM: RAND/UCLA Appropriateness Method

Edited by Javad Sarvestan; peer-reviewed by Rebecca H Correia; submitted 18.May.2026; final revised version received 16.Jul.2026; accepted 22.Jul.2026; published 07.Aug.2026

Please cite as:

Tavares S, Beaney T, Reisberg A, Henley A, Edwards A, Belsey D, Downey L

Codeveloping a Set of Quality of Care Indicators for People Living With Multiple Long-Term Conditions in Primary Care: Protocol for a Modified RAND/UCLA Study

JMIR Res Protoc 2026;15:e99217

URL: <https://www.researchprotocols.org/2026/1/e99217>

doi: [10.2196/99217](https://doi.org/10.2196/99217)

© Sara Tavares, Thomas Beaney, Arad Reisberg, Ania Henley, Angela Edwards, David Belsey, Laura Downey. Originally published in JMIR Research Protocols (<https://www.researchprotocols.org>), 07.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.researchprotocols.org>, as well as this copyright and license information must be included.