OPEN ACCESS

International survey of people living with chronic conditions: development and evaluation of the PaRIS Patient Questionnaire (PaRIS-PQ) in 18 countries

Jose M Valderas ,[1,2,3] Ian Porter ,[1] Jonathan P Evans,[1] Monique Heijmans,[4] Mieke Rijken ,[4,5] Oliver Groene,[6,7] Janika Bloemeke-Cammin,[6] Rosa Sunol,[8,9,10] Rachel Williams,[11] Marta Ballester,[8,9,10] Katherine de Bienassis,[12] Candan Kendir,[12,13] Frederico Guanais,[12] Dolf de Boer,[4] Michael van den Berg[12]

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/bmjqs-2025-018548).

For numbered affiliations see end of article.

Correspondence to
Professor Jose M Valderas; j.m.valderas@exeter.ac.uk

JMV and IP are joint first authors.

Received 10 February 2025
Accepted 5 August 2025

Check for updates

© Author(s) (or their employer(s)) 2025. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ Group.

To cite: Valderas JM, Porter I, Evans JP, *et al*. *BMJ Qual Saf* Epub ahead of print: [please include Day Month Year]. doi:10.1136/bmjqs-2025-018548

ABSTRACT

Background The Patient Reported Indicator Surveys (PaRIS) initiative of the Organisation for Economic Co-operation and Development aimed to develop a valid and reliable instrument for self-reported assessment of health outcomes and experiences in primary and ambulatory care for people living with chronic conditions. This paper reports on the development and evaluation of the PaRIS Patient Questionnaire (PaRIS-PQ) in an 18-country field trial.

Methods Following systematic reviews of instruments measuring core domains in the PaRIS conceptual framework (outcomes, experiences, health and healthcare capabilities and health behaviours), four instruments for each domain were shortlisted, and their psychometric performance was assessed using the Evaluating the Measurement of Patient-Reported Outcomes tool. In a modified Delphi procedure, one instrument was selected for each domain, along with additional relevant items. The preliminary English questionnaire was translated and cognitively tested. Psychometric evaluation was conducted on field trial data at item (missingness, distribution, ceiling/floor effects) and scale level (reliability, structure, construct validity).

Findings 217 instruments were identified measuring one or more of the domains of the conceptual framework. Following mapping and prioritisation, the first draft of the PaRIS-PQ included 118 items. In the field trial (18 countries, 10 894 patients) median completion time ranged 24–33 minutes. PaRIS-PQ performed well at item and scale level. Reliability was ≥0.70 for most relevant measures at patient level, but not at practice and country level. Validity was adequate overall. Removal of additional items (3) resulted in the final PaRIS-PQ (115 items).

Interpretation PaRIS-PQ demonstrates adequate psychometric performance for measuring experiences and outcomes of primary care for people living with chronic conditions. The instrument facilitates the collection of essential information for health policy and systems decision-making.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ There are a number of patient questionnaires measuring either outcomes, experiences or health and healthcare capabilities and health behaviours. However, these instruments primarily focus on each domain in isolation and do not provide a comprehensive picture of primary care performance for people with chronic conditions.

BACKGROUND

Countries around the world are redesigning their health systems to respond to the challenge of growing populations living with chronic conditions.[1 2] Primary care has a vital role to play in the prevention, identification and management of chronic conditions through its core functions of first access, coordination, continuity and comprehensiveness.[3 4] The focus on person-centredness in primary care, is also fundamental for orienting healthcare (including self-management) for patients living with chronic conditions.[5–7]

Using appropriate indicators rooted in what matters to patients, performance assessment and feedback can help primary care and health systems ensure person-centred management of chronic conditions. For example, patient-reported experiences and outcomes of care, which are typically captured using standardised

WHAT THIS STUDY ADDS

⇒ The scope of this questionnaire is more comprehensive than those previously developed and is unique in capturing what is important for patients diagnosed with chronic conditions in primary care. It is based on a multilevel conceptual framework, linking the care experiences and outcomes reported by individual persons using primary care to the characteristics of their primary care providers, to the characteristics of the health systems in which the providers operate.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ It will enable cross-country comparisons of people's outcomes and experiences of care, helping policymakers identify, disseminate and implement best practices. This will fuel international learning and foster a dialogue with patients and service providers on how to further improve the performance and people-centredness of primary care services.

patient-reported experience measures (PREMs)[8] and patient-reported outcome measures (PROMs),[9] aim to assess quality (including safety) and outcomes from the patient perspective.

In line with these principles, the Health Ministers of Organisation for Economic Co-operation and Development (OECD) Member countries mandated the OECD to lead an effort to develop comparable cross-country indicators for patients' experiences of healthcare and their self-reported health outcomes.[10 11] This led to the OECD's Patient-Reported Indicator Surveys (PaRIS) initiative, which aims to fill the critical gap in a measurement framework that includes patient perceptions of access to healthcare, quality of life, pain, physical functioning and psychological well-being. The PaRIS initiative set out to develop, standardise and implement a new generation of generic (ie, non-disease specific) patient-reported indicators that measure the outcomes and experiences of healthcare that matter most to adults with chronic conditions using primary or (generalist) ambulatory care services in participating countries.[12]

A key component of this project has been the development of the PaRIS Patient Questionnaire (PaRIS-PQ). This first-of-its-kind international survey provides a unique set of indicators on how people living with chronic conditions experience healthcare (its quality and safety) and how this impacts their lives. The PaRIS-PQ operationalises the domains identified in the PaRIS conceptual framework[13]: patient-reported outcomes; patient-reported experiences of care (including patient safety and overall perceived quality of care); health and care capabilities; health behaviours, and sociodemographic characteristics.[13] The framework, developed through a systematic, accountable and inclusive process, also considers delivery system characteristics (clinic and main healthcare professional, captured in the PaRIS Primary Care Practice Questionnaire (PaRIS-PCPQ)[14]) and health system, policy and context.

The PaRIS-PQ aims to address the following questions[12]:

- What are the patient-reported outcomes of primary care service users aged 45 and over with chronic conditions, compared with those without chronic conditions, in the areas of symptoms, physical, mental and social functioning, self-reported health and health-related quality of life? How do these results vary across countries?
- What are the experiences of primary care service users aged 45 and over with chronic conditions, compared with those without chronic conditions, in the areas of access, comprehensiveness, continuity, coordination, safety and people-centredness of care, self-management support, trust and overall perceived quality of care? How do these results vary across countries?
- How do patient-reported outcomes and care experiences vary for primary care service users aged 45 and over with chronic conditions by background characteristics such as age group, gender, education level, occupational status, household composition, health-risk behaviours, level of multimorbidity, disease status and confidence in managing one's own care?

Information provided by patients on the PaRIS-PQ is complemented by providers' responses to the PaRIS Primary Care Practice Questionnaire (PaRIS-PCPQ),[14] which was developed in tandem based on the same conceptual framework. Linking patient self-reports with provider-reported assessments facilitates the evaluation of how the organisation of primary care practices and their care processes impact care experiences and outcomes.[14] The PaRIS-PCPQ is designed to be completed online by medical or non-medical staff working in a primary care practice, providing information on practice profile, care coordination, chronic care management, patient follow-up and respondent characteristics. The final version (based on data from 540 primary care practices across 17 countries) contains 34 items and has a median completion time of 20 minutes.

The PaRIS-PQ aims to provide information on the patient-reported outcomes and experiences of care of adults 45 years and over with chronic conditions. The main purpose is to provide estimates of health system performance, particularly in relation to care for chronic conditions in primary care and ambulatory settings, facilitating comparative performance assessment between and within countries. The nested design of the survey (practices identified at random within each country and patients identified at random within each participating practice), as well as the patient-centric nature of measurement, overcomes some of the limitations of existing initiatives using patient-reported information for international comparative purposes.

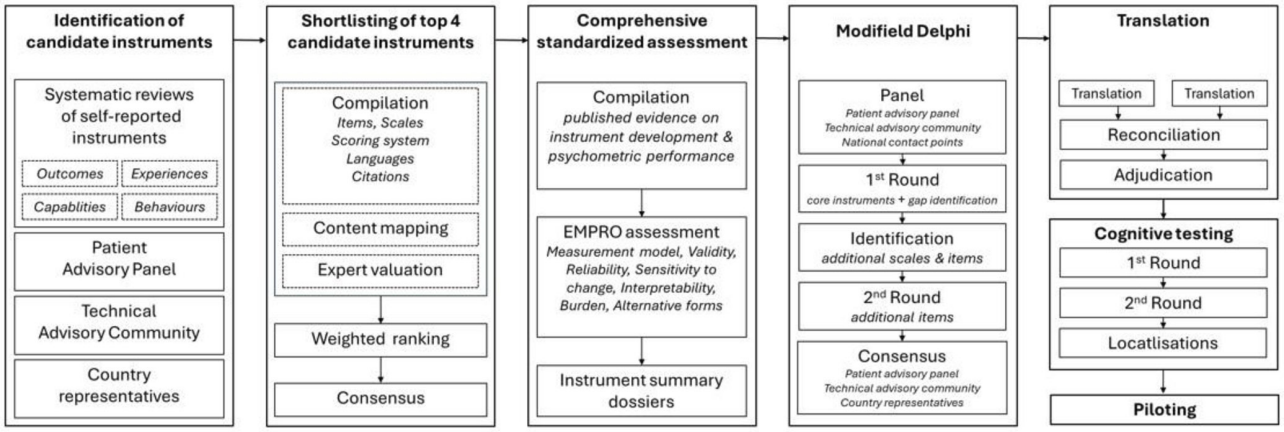


Figure 1 Development process of the PaRIS Patient Questionnaire

The objective of this paper is to report on (1) the development of the PaRIS-PQ, including cross-cultural translation and cognitive testing and (2) the field trial piloting of the PaRIS-PQ across 18 countries, including an evaluation of psychometric properties.

METHODS

The development of PaRIS-PQ followed a published protocol, which, outlined collaborative methods, including reliance on existing suitable measurement instruments, to ensure that the resulting questionnaire would have the best possible psychometric performance in terms of validity (including cross-cultural validity) and reliability.[12] Multiple guidelines of varying quality have been proposed for the development of patient-reported measures[15] and the project was conducted in accordance with recommended standards proposed by the scientific advisory task force of the International Society for Quality of Life Research.[16] The development process is summarised in figure 1.

Conceptual framework

The development of the PaRIS-PQ was underpinned by a conceptual framework, produced through narrative review and the participation of international stakeholders, including people over 45 with lived experience of chronic conditions, and a Patient Advisory Panel.[13] This framework incorporates four broad domains.

- Health behaviours (HB).
- Health and healthcare capabilities (HHC) (knowledge, skills and confidence of individuals for living well and managing their health concerning their interactions with the healthcare system).
- Patient-reported experiences (individual perceptions about the range of interactions that patients have with the healthcare system or professionals on the degree to which they respond to their needs).
- Patient-reported outcomes (any aspect of a patient's health status that comes directly from the patient).[12 17]

Each of these core domains is influenced by individual and sociodemographic factors (table 1):

Identification, shortlisting and evaluation of candidate core instruments

First, we identified the best candidate instrument for comprehensive measurement of each domain (*core instruments*). In doing so, it was important to select validated instruments appropriate for adults across a wide range of countries and contexts, measuring patient-reported outcomes, experiences of care, health and HHC and HBs and lifestyle. We conducted a series of bespoke literature reviews to identify candidate instruments (online supplemental appendix 1). An initial long list of eligible instruments for consideration for inclusion was produced and supplemented with feedback from PaRIS Patient Advisory Panel (PaRIS-PP), Technical Advisory Community (TAC) and the OECD Secretariat, more details on which are provided in the project protocol.[12]

Two experts in patient-reported measurement (JMV and IP) extracted key information for each instrument's characteristics, including name, number of items, number of scales (their name and what they assess), scoring system and available languages. They mapped the candidate instruments onto the domains and subdomains in the framework. Instruments were classified as measuring each domain and subdomain in the framework based on both the corresponding conceptualisation and definition of each domain and subdomain as well as the questionnaire scales, scoring systems, items and available supporting documents identified in the previous stages. Each instrument, therefore, could measure multiple domains and subdomains.

A single core instrument was selected for each of the four domains. The top four candidate instruments for each domain were shortlisted using a pragmatic weighted ranking (table 2). Ranking criteria and weights reflected priorities for shortlisting: maximising

Table 1 The patient-centred domains, subdomains and variables in the PaRIS conceptual framework of experiences and outcomes of people living with chronic conditions, managed in primary and ambulatory care

Core domains	Subdomains	Items
Individual and sociodemographic factors	Demographic, biometric and morbidity factors	▶ Age ▶ Sex ▶ Body mass index ▶ Chronic conditions ▶ Burden of morbidity ▶ Disability
	Socioeconomic factors	▶ Ethnicity ▶ Sexual orientation ▶ Education level ▶ Occupational status ▶ Income ▶ Migrant status ▶ Urbanisation ▶ Household composition and size ▶ Social support
Health behaviours and lifestyle	Physical activity	▶ Physical activity
	Diet	▶ Diet
	Tobacco use	▶ Tobacco use
	Alcohol use	▶ Alcohol use
Health and healthcare capabilities	Health and healthcare capabilities	▶ Health and healthcare capabilities
Patient-reported experiences of care	Access	▶ Access
	Comprehensiveness	▶ Comprehensiveness
	Continuity	▶ Continuity
	Coordination	▶ Coordination
	People-centred care	▶ Person-centred care ▶ Decision-making ▶ Interacting with health professionals
	Safety	▶ Safety
	Self-management support	▶ Self-management support
	Trust	▶ Trust
	Overall perceived quality of care	▶ Overall perceived quality of care
Patient-reported outcomes	Symptoms	▶ Pain ▶ Fatigue ▶ Shortness of breath ▶ Symptoms of depression ▶ Symptoms of anxiety
	Functioning	▶ Physical functioning ▶ Mental functioning ▶ Social functioning
	Self-reported health	▶ Overall self-reported health status
	Health-related quality of life	▶ Overall health-related quality of life

PaRIS, Patient Reported Indicator Surveys.

instrument coherence and perceived quality and minimising administration burden. These weights were piloted with a subsample to ensure that the proposed weighted ranking elicited ranking variations in the expected direction and of the expected magnitude. A standardised evaluation of the metric properties and

Table 2 Ranking criteria for candidate core instruments

Criterion	Rationale	Weight
Number of subdomains the instrument had been mapped onto	To prioritise comprehensive instruments in order to minimise the number of additions necessary while maximising consistency in formatting	0.4
Number of items	To minimise administrative burden	0.2
Expert judgement	To consider overall expert perception of the quality and relevance of the instrument	0.2
Citations in Google Scholar to the first publication	To provide an objective yet indirect estimate of the quality and relevance of the instrument	0.1
Number of languages the instrument was available	To further provide evidence of international applicability and minimise the subsequent burden of translation and adaptation	0.1

ease of use of each instrument was then independently conducted in pairs (combinations of JMV, JPE and IP) using the Evaluating the Measurement of Patient-Reported Outcomes (EMPRO) tool[18] based on systematic searches for core references for each candidate instrument. The following attributes were assessed: measurement model, reliability, validity, responsiveness, interpretability, burden and alternative forms. For each candidate instrument, seven attribute-specific scores were generated, and one overall non-weighted EMPRO score was calculated (range 0–100, worst to best).

Modified Delphi for the selection of the core instruments

A modified Delphi process, a methodology widely used in health and social science research,[19] was conducted. The modified approach differs from the traditional Delphi approach in that it begins with an initial list of items, rather than the expert panel generating them themselves. Members of the PaRIS-PP, TAC and National Project Managers (NPMs), the project managers who lead development and implementation in participating countries, participated via an online Delphi portal. This ensured we had appropriate expertise encompassing patient lived experience of chronic conditions, patient-reported measurement, primary care, chronic disease management and international representation to provide a regional context. A dossier was compiled for each candidate instrument to support decision-making, including relevant evidence extracted from documents as identified in the previous steps. Participants were invited to rank all the candidate instruments from their first to last (fourth) choice for each broad domain, considering specific characteristics such as comprehensiveness, psychometric robustness, administration burden, international comparability and overall suitability. The a priori criteria that signalled a lack of consensus, and

the need for subsequent rounds to identify the core set of instruments, were as follows:

- Multiple instruments with a median (Md) ranking of 1 (first choice).
- Significant variation in rankings for the top-ranked instrument (IQR>1).
- More than one-third of participants ranking the second-best instrument higher than the top-ranked one.

To ensure the comprehensiveness of the PaRIS-PQ instrument, additional measures (either full instruments, scales, item sets or single items) were identified by the team (IP, JMV) for those domains that were not adequately covered (either partially or in their entirety) by the core instrument and based on the pool of instruments and items identified in the previous steps. Subsequent Delphi rounds aimed to gather consensus for obtaining the final set of items and scales. Nearly all of the questions were taken from validated sources, with the exception of a limited number developed ad hoc to elicit necessary information relating to aspects such as the purpose of recent consultations and healthcare professionals seen.

A draft version of the PaRIS-PQ and an accompanying report outlining the development process was submitted to key stakeholders: the PaRIS-PP, TAC, NPMs and the Working Party for PaRIS (the official OECD body in which participating countries are represented). Multiple iterations were produced based on feedback received, and the process culminated when no further changes were deemed necessary.

Translation and cognitive testing

The source English versions of the patient questionnaire were first translated into a French reference questionnaire and then into the main languages of each participating country (18 languages across 22 countries) (online supplemental appendix 2) by cApStAn, a linguistic quality control agency, in collaboration with local translators and based on the Translation, Review, Adjudication, Pre-testing and Documentation approach, the established leading practice in international surveys.[20] A collaborative and iterative approach prioritised that translated versions reflected the local context while maintaining the integrity of the source.

The cognitive testing process was overseen by the multinational market research and public opinion organisation Ipsos. Ipsos provided training for NPMs in each country, along with a cognitive testing manual and supporting documentation. Experienced researchers in the participating countries received additional in-depth training before conducting two main rounds of cognitive testing with up to 20 participants per country. Consistent with the design of the PaRIS-PQ, participants were required to be aged 45 years or older at the time of sampling, living in a private household in the community, and to have had at least one registered contact with a primary care provider in the preceding 6 months. Interviews were conducted face-to-face or remotely (via telephone/online teleconference) using online or paper-based questionnaires. Each country provided a structured analysis grid to ensure comparable evidence was available to support the need for any changes, and each issue was logged and categorised using the Cross National Error Source Typology framework.[21] The typology covered the following issues: source questionnaire design, cultural portability (eg, applying concepts across different cultural contexts), translation (eg, finding an equivalent term in a specific language), where minor or no changes were needed. All logged issues, along with their supporting evidence, were reviewed to assess and implement necessary changes to both the source and localised versions of the PaRIS-PQ. This thorough review process resulted in a revised and robust questionnaire, well-suited for diverse languages and contexts.

Piloting and psychometric evaluation

Piloting of the PARIS-PQ was conducted in 18 countries: Australia, Belgium, Canada, Czechia, France, Greece, Iceland, Italy, Luxembourg, Netherlands, Norway, Portugal, Romania, Saudi Arabia, Slovenia, Spain and the United Kingdom (England and Wales contributed separately). Data collection was administered through a combination of online, telephone and paper questionnaires, taking place between March and November 2022. Ethics approval was obtained according to the process and requirements of each participating country.

Interested countries declared participation with the OECD Health Secretariat. As a country with a specific action plan with the OECD, Saudi Arabia also participated in the PaRIS initiative.[22] NPMs appointed in each country were provided with detailed guidelines on sampling and recruitment. Primary care practices were considered to be eligible if they were staffed with care professionals licensed to serve the general population of a community and provide ambulatory generalist medical care (ie, in an outpatient setting), including services addressing chronic care management. Eligible criteria applied for patients were as follows: (1) to be aged 45 years or older at the time of sampling; (2) living in the community (in a private household, ie, not in a long-term care facility, healthcare or other residential institution); and (3) having had at least one registered contact with a primary care provider—either face-to-face, by telephone or online—for any medical or administrative reason, during the 6 months preceding the selection procedure.

NPMs constructed a sampling frame for the Field Trial for drawing a probability sample of eligible primary care practices and patients (national representativeness was not required for the Field Trial). NPMs were instructed to draw samples that were sufficiently large to result in 25 participating primary care practices and 50 participating patients per practice. These

target numbers were based on statistical simulations carried out on a large dataset of patient-reported care experiences from 34 countries.[23] The simulations aimed to optimise the balance between the reliability of patient-reported indicators and the number of responding primary care practices and service users. For smaller countries, a lower target for practices was set (10).

A psychometric evaluation was conducted at both the item level (missingness, response distribution, ceiling/floor effects) and the scale level (reliability, structure and validity). Cronbach's alpha, a measure of reliability, was calculated for each of the (sub) scales of the included instruments. We also constructed multilevel models for estimating reliability coefficients at patient, primary care practice and country levels. The reliability of the constructs at these higher levels is important to enable multilevel regression analyses in answering the primary research questions of the PaRIS-PQ.[12] Reliability coefficients ≥0.70 for all levels were considered to indicate good reliability.

For already extensively validated scales, we conducted confirmatory factor analysis (CFA) and exploratory factor analysis (EFA) for all other scales. We also estimated a series of models to get insight into the extent to which the factor structure remained invariant across countries (multigroup CFA): (a) a model assuming the same factor structure across countries, (b) a more restrictive model also assuming that the factor loadings were similar across countries ('metric invariance'), (c) the most restrictive model assuming that the factor structure, the factor loadings and the error terms were similar across countries ('scalar variance').

Concerning the evaluation of construct validity, how well the PaRIS-PQ assessed the constructs it was designed to measure, we hypothesised the various PROMs to be moderately (0.30–0.50) to strongly (>0.50) correlated to each other. Similarly, this was expected for the multiple PREMs and HHC. We expected significant yet less strong associations regarding patients' HBs. We expected PROMs and PREMs to correlate weakly (0.10–0.30) or moderately (0.30–0.50). We expected weak correlations between health/risk behaviours and PROMs, as well as between health and HHC and PROMS, reasoning that many other factors (eg, genetic, demographic, socioeconomic, morbidity) also impact patients' reported health outcomes, beyond health/risk behaviours and capabilities. PREMs were expected to correlate more strongly with patients' health and care capabilities than with their health and risk behaviours, based on the assumption that patients prefer and are capable of playing an active role in managing their health and care, which will impact their care experiences. In contrast, a more negligible impact on patients' care experiences can be expected from their health/risk behaviours. Finally, patients' health and care capabilities were expected to correlate weakly with health/risk behaviours.

Final version of the PaRIS-PQ

Based on issues previously identified in the cognitive evaluation, feedback from countries and evaluation of field trial data, a revised version of the patient questionnaire was created.

RESULTS

Overall, 217 instruments were identified that measured one or more of the domains of the conceptual framework (21 HB, 62 HHC, 63 PREMs, 71 PROMs). The shortlisting process resulted in four instruments, each measuring a broad domain in the conceptual framework (online supplemental appendix 3). A consensus was reached in the EMPRO evaluation, with minor disagreements resolved quickly through discussion.

19 participants took part in the Delphi procedure. The first stage resulted in the confirmation of a comprehensive core set of instruments as the basis for measuring all domains and the majority of subdomains in the framework, followed by a subsequent stage that confirmed additional instruments covering areas not adequately covered. The draft questionnaire was structured into four main sections ('Your health', 'Managing your health and health care', 'Your experience of health care' and 'About yourself'). It included four existing instruments (PROMIS Scale V.1.2—Global Health,[24] The World Health Organization-Five Well-Being Index (WHO-5),[25 26] Porter-Novelli Scale[27] and the person-centred coordinated care experience questionnaire (P3CEQ)[28]) and a total of 118 items.

Cognitive testing was conducted in 18 languages across 22 countries, including five English adaptations, with an average of 17 participants per country. 845 points of feedback were logged and collated, covering issues relating to the source instrument, cultural portability, technical translation difficulties and other comments without specific recommendations. This resulted in four items in the source questionnaire being replaced with alternatives. 11 minor modifications were also made: clarifying explanations prior to questions (2), question stem wording (6) and response options (3).

See figure 2 for items and their sources mapped onto the conceptual framework.

Psychometric evaluation

The field trial psychometric evaluation was based on 11 999 patients who provided consent across 18 countries, of which 10 894 completed the patient questionnaire (91%). 13 out of 15 countries offered the questionnaire online. Six countries offered only online administration of the questionnaire. Seven countries also offered paper questionnaires. Two countries, Saudi Arabia and Spain, offered the patient questionnaire in a telephone mode, as this better fitted the preferences and values of their citizens (see online supplemental figure 1).

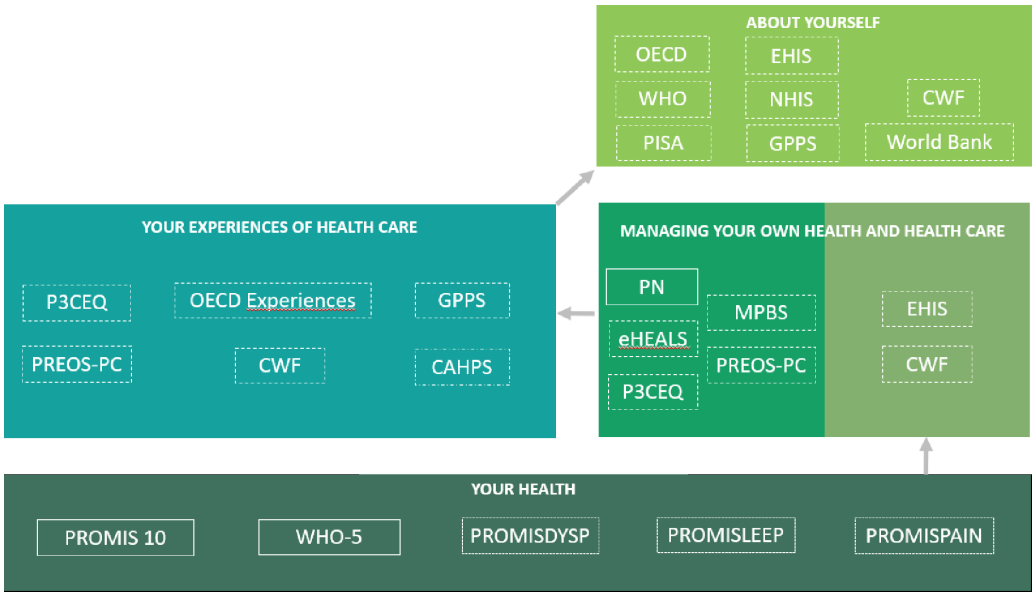


Figure 2 Source of items in the PaRIS-Patient Questionnaire. CAHPS, Consumer Assessment of Healthcare Providers and Systems Surveys; CWF, Commonwealth Fund Survey; eHEALS, The eHealth Literacy Scale; EHIS, European Health Interview Survey; GPPS, GP Patient Survey; MPBS, Medicare Patient Beneficiary Survey; NHIS, National Health Interview Survey; OECD, various OECD surveys; PaRIS, Patient Reported Indicator Surveys; P3CEQ, Person-Centred Coordinated Care Experience Questionnaire; PISA, Programme for International Student Assessment; PN, Porter-Novelli Consumer Preferences Scale; PREOS-PC, Patient Reported Experiences and Outcomes of Safety in Primary Care; PROMISDYSP, PROMIS Item Bank—Dyspnea; PROMISLEEP, PROMIS Item Bank—Sleep Disturbance; PROMISPAIN, PROMIS Adult Short Form—Pain Interference; PROMIS-10, PROMIS Scale V.1.2—Global Health; WHO-5, WHO Well-Being Index; World Bank, World Bank Model-Disability-Survey.

The median time to complete was 24 to 33 minutes across countries. Dropout rates were generally low; only three questions (out of 118) had a high rate of missingness (more than 10% of answers missing). Answers to most questions did not point to any ceiling or floor effects (see online supplemental appendix 4). Many questions had a fairly normal, bell-shaped distribution of valid answers, with a few exceptions, such as question 12 (*How would you rate your pain on average?*).

All (multiple-item) PROMs showed good reliability at the patient level (both in the single-level model and the multi-level model), except for proposed scales for Fatigue and Sleep (0.56 and 0.67, respectively).

Table 3 Reliability of multi-item patient-reported outcome measures

		Single-level reliability	Multilevel reliability		
Scale	**Items**	**Cronbach's α**	**Patient level**	**Practice level**	**Country level**
Physical functioning	PROMIS Scale Global Health—Physical Health Summary Score (4 items)	0.77	0.76	0.57	0.60
Mental functioning	PROMIS Scale Global Health—Mental Health Summary Score (4 items)	0.80	0.79	0.52	0.77
Well-being	WHO Well-Being Index (5 items)	0.89	0.88	0.47	0.55
Pain (new)	PROMIS Adult Short Form V.1.0—Pain Interference question and PROMIS Scale V.1.2—Global Health question (2 items)	0.72	0.87	0.54	0.51
Fatigue (new)	PROMIS Scale V.1.2—Global Health question and WHO Well-Being Index (WHO-5) question (2 items)	0.70	0.56	0.33	0.61
Sleep (new)	PROMIS Item Bank V.1.0—Sleep Disturbance question and WHO Well-Being Index (WHO-5) question (2 items)	0.67	0.67	0.36	0.54
Symptoms of depression (new)	PROMIS Scale V.1.2—Global Health question and two questions from the WHO Well-Being Index (WHO-5) (three items)	0.78	0.77	0.45	0.65
Symptoms of anxiety (new)	PROMIS Scale V.1.2—Global Health question and WHO Well-Being Index (WHO-5) question (2 items)	0.73	0.72	0.35	0.72
Social functioning (new)	Two questions from PROMIS Scale Global Health (2 items)	0.80	0.77	0.39	0.70

Overall, no PROMs could be reliably assessed at practice or country level (see table 3).

PROMs showed an acceptable fit when the data from all countries were taken together for standardised factor loadings: The PROMIS Scale V.1.2 – Global Health—*Physical scale* (0.65–0.71.), PROMIS Scale V.1.2—Global Health—*Mental scale* (0.58–0.83), *WHO-5 Wellbeing Index* (0.75–0.83). The multigroup confirmatory factor analyses demonstrated that the factor structure held well across all countries (configural invariance). The metric model exhibited a more substantial misfit as the SRMR (Standardized Root Mean Square Residual) increased substantially, indicating that factor loadings differed across countries. The more restrictive scalar model could not be held.

Reliability coefficients ≥0.70 for all levels were considered to indicate good reliability.

The P3CEQ total score demonstrated good internal consistency reliability at the patient level, both in single-level and multilevel analyses. This also held for the person-centredness subscale. The care coordination scale at the patient level in the multilevel analysis was lower than 0.70 (see online supplemental table 1). PREMs correlated weakly to moderately with each other, except for *person-centred care* and *care coordination* (0.82), *person-centred coordinated care experience* and *care coordination* (0.92) and *person-centred coordinated care experience* and *person-centred care* (0.96). Initial EFA yielded only one factor with an eigenvalue greater than 1 (2.77); the second-largest factor had an eigenvalue of 0.24. This suggests that the P3CEQ assessed a single unidimensional concept of person-centred coordinated care experiences rather than two distinct but related concepts: experience of person-centredness and experience of coordination of care.

Two dimensions of the Porter-Novelli Scale showed acceptable reliability in single-level analysis. The patients' confidence in the self-management construct showed good internal consistency and reliability in single-level analysis. The EFA of the Medicare Current Beneficiary Survey item set, combined with the question from the eHealth Literacy Scale, resulted in one factor with an eigenvalue larger than 1 (2.75). The factor loadings of all questions are high (above 0.60), with only one lower loading (0.43), which was still considered acceptable. This analysis confirmed the construction of a single scale to assess patients' confidence in self-management.

As per our a priori hypotheses, all PROMs correlated strongly with each other (range: 0.47–0.82) and correlated more weakly with PREMs (range: −0.27–0.28), HB (range: −0.39–0.28) and HHC (range: −0.36–0.36) than anticipated. PREMs correlated weakly to moderately with each other, except for *person-centred care* and *care coordination* (0.82), *person-centred coordinated care experience* and *care coordination* (0.92) and *person-centred coordinated care experience* and *person-centred care* (0.96). The other correlations between PREMs ranged between −0.09 and 0.72. As expected, the correlations between PREMs and HB were weak (range: −0.25–0.16) and stronger between PREMs and HHC (range: −0.27–0.32). HHC was correlated more weakly with HB than anticipated (−0.11–0.27).

Final changes to PaRIS-PQ

Based on issues previously identified in the cognitive evaluation, feedback from countries and evaluation of field trial data, a revised version of the patient questionnaire was created. A summary of the changes is outlined in online supplemental appendix 5, and a list of abbreviations is included in online supplemental appendix 6. The final source version of the postfield trial PaRIS-PQ is available online.[29]

DISCUSSION

PaRIS-PQ is a comprehensive, first-of-its-kind, international questionnaire developed based on a conceptual framework for people living with chronic conditions, following an inclusive approach. International cognitive testing helped improve the quality and provide evidence for the face validity of the questionnaire, which was grounded in patients' experiences. The questionnaire offers an invaluable perspective on outcomes and experiences among the largest and fastest-growing group of healthcare users: people with chronic conditions, with the potential to inform policymakers about areas of performance needing improvement in their country.

Evaluation of field trial data provided evidence of the questionnaire's performance at both the item and scale levels. The questionnaire length was not considered problematic, and dropout rates were generally low; refinements were made by removing or substituting some questions. Measures showed good reliability at the patient level, both in the single-level and multilevel models; however, this was not the case at the practice and country levels. Multi-group confirmatory factor analyses demonstrated that the factor structure held well across all countries (configural invariance), but more so for the more restrictive metric and scalar variance models.

Although other large-scale international health surveys exist, such as the Commonwealth Fund[30 31] and People's Voice Surveys,[32–34] the PARIS-PQ is unique in linking patients with their regular primary care practice, emphasising health systems performance in the delivery of primary care for chronic care management and focusing on patient-reported outcomes. This domain has merited limited attention in the other initiatives. Also, the scope of the conceptual framework on which the PaRIS-PQ is based is more extensive than those previously developed, reflecting the multilevel nature of the causal associations between constructs, providing a basis for hypotheses in relation

to associations between variables,[35] as detailed in the PaRIS data analysis plan based on this framework.

There are multiple strengths in the development and evaluation of the PaRIS-PQ. We followed a replicable systematic process conducted over multiple rounds using innovative, state-of-the-art methods. Key stakeholders and experts, including the Working Party for PaRIS, PaRIS-PP and TAC, provided regular input to ensure that the questionnaire has been developed to a high standard and is relevant to what is important to patients.

We also acknowledge that there have been limitations in developing the PaRIS-PQ. A challenge of an international undertaking such as this is balancing the inherent tension between comprehensiveness, accuracy and feasibility of administration as well as adaptation to national priorities. It is desirable but challenging to create a questionnaire that captures what is important to patients with chronic conditions while maintaining brevity. In doing so, we have faced the common minimum denominator problem, settling for solutions that achieved consensus, allowing for comparability, but perhaps at the expense of specificity for each country. The large number of domains and subdomains requires a minimum number of questions for any given construct. However, we acknowledge that some areas of the conceptual framework can benefit from enhanced coverage, such as comprehensiveness of care. The PaRIS-PQ was mainly based on previously validated questionnaires, which had advantages but also meant that we relied on other sources of information developed in different contexts.

A low reliability at the practice level implies that no firm conclusions can be drawn about differences between practices (within countries) in their patients' experiences of symptoms, functioning or health-related quality of life, at least not with the low total number of practices participating in the Field Trial. Regarding the reliability of the constructs assessed at the country level, the results are better, although still insufficient to draw conclusions about differences between countries. Evaluation of psychometric properties is a cumulative process. It is essential to note that this is the pilot implementation of the PaRIS-PQ, and it is anticipated that future iterations with a higher number of patients per practice will likely yield higher reliability coefficients.

PROM correlations with PREMs, HB and HHC were in the same direction as hypothesised, and some were weaker than anticipated. Discrepancies may be due to our hypotheses not being correct in the first place. Our predictions may have overestimated associations between measurements that are of a generic nature rather than disease and medical service-specific. The latter may well be the ones more aligned with our hypothesis. Lower than expected reliability may also have contributed to lower correlation estimates. The body of evidence on the validity of the standardised measures, as assessed with EMPRO, supports our confidence in the validity of the measures.

In conclusion, we present the inclusive, systematic and replicable development of the PaRIS-PQ and provide evidence of cross-cultural validity as well as a preliminary evaluation of its psychometric properties. It is anticipated that future large-scale implementations will provide additional evidence of the instrument's reliability, structure and validity. The instrument is now freely available and has the potential to provide valuable insights into how people living with chronic conditions experience healthcare, ultimately shaping healthcare systems to be more responsive to their needs.

Author affiliations
[1]Department of Health & Community Sciences, Exeter Collaboration for Academic Primary Care (APEx), University of Exeter, Exeter, UK
[2]Centre for Research in Health Systems Performance (CRiHSP), National University of Singapore, Singapore
[3]Department of Family Medicine, National University Health System, Singapore
[4]Netherlands Institute for Health Services Research (NIVEL), Utrecht, The Netherlands
[5]Department of Health and Social Management, University of Eastern Finland, Joensuu, Finland
[6]Optimedis AG, Hamburg, Germany
[7]Faculty of Management, Economics and Society, University of Witten/Herdecke, Witten, Germany
[8]Avedis Donabedian Research Institute (FAD), Barcelona, Spain
[9]Network for Research on Chronicity, Primary Care, and Health Promotion (RICAPPS), Barcelona, Spain
[10]Universitat Autònoma de Barcelona, Barcelona, Spain
[11]Ipsos, London, UK
[12]Directorate for Employment, Labour and Social Affairs, Organisation for Economic Co-operation and Development (OECD), Paris, France
[13]Department of Public and Occupational Health, Amsterdam UMC location Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

Acknowledgements PaRIS is an OECD initiative. An international consortium, consisting of Nivel, Exeter University, Ipsos, Avis Donabedian Research Institute, and Optimedis, led the development and implementation of the PaRIS-PQ. The authors thank the Working Party for PaRIS, the members of the PaRIS Patient Advisory Panel, the Technical Advisory Community, and various experts for their roles throughout the study. The authors also thank the national project management teams and participating primary care professionals and patients.

Contributors JMV and IP were involved in every stage of the project and contributed equally to the paper as joint first authors. JMV is the guarantor. JPE assessed the psychometric performance of shortlisted instruments using the Evaluating the Measurement of Patient-Reported Outcomes (EMPRO) tool. MR, MH, OG, JB, RS, RW, MB and DdB all contributed as members of the PaRIS Consortium and were involved in every stage in developing the patient questionnaire and drafting the paper. KdB, CK, FG, and MvdB represented the OECD in providing oversight in each stage of the project and contributed to the drafting of the paper.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval In each participating country, the National Project Manager coordinated the processes for obtaining relevant legal and ethical approvals, as appropriate to their jurisdictions. For more information on the approval process

in general or the specific requirements of individual countries, please contact the OECD Secretariat. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer-reviewed.

Data availability statement Data are available upon reasonable request.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iDs
Jose M Valderas http://orcid.org/0000-0002-9299-1555
Ian Porter http://orcid.org/0000-0002-2440-8727
Mieke Rijken http://orcid.org/0000-0001-6070-4091

REFERENCES

1 Mirzoev T, Kane S. What is health systems responsiveness? Review of existing knowledge and proposed conceptual framework. *BMJ Glob Health* 2017;2:e000486.
2 World Health Organisation. A vision for primary health care in the 21st century. Geneva, 2018.
3 World Health Organisation. A vision for primary health care in the 21st century: towards universal health coverage and the sustainable development goals. 2018.
4 Nolte E, Pitchforth E. Assessing chronic disease management in european health systems. Nolte E, Knai C, Saltman R, eds. Copenhagen: WHO Regional Office for Europe, 2014.
5 Mead N, Bower P. Patient-centredness: a conceptual framework and review of the empirical literature. *Soc Sci Med* 2000;51:1087–110.
6 Valderas JM, Gangannagaripalli J, Nolte E, *et al*. Quality of care assessment for people with multimorbidity. *J Intern Med* 2019;285:289–300.
7 Nolte E, McKee M. Caring for people with chronic conditions a health system perspective. Nolte E, McKee M, eds. Maidenhead: Open University Press, 2008.
8 Jamieson Gilmore K, Corazza I, Coletta L, *et al*. The uses of Patient Reported Experience Measures in health systems: A systematic narrative review. *Health Policy* 2023;128:1–10.
9 Valderas JM, Alonso J. Patient reported outcome measures: a model-based classification system for research and clinical practice. *Qual Life Res* 2008;17:1125–35.
10 Slawomirski L, Berg M, Karmakar-Hore S. Patient-Reported indicator survey (Paris): aligning practice and policy for better health outcomes. *World Med J* 2018;64:8–14.
11 OECD. Recommendations to oecd ministers of health from the high-level reflection group on the future of health statistics. Available: https://www.oecd.org/en/publications/recommendations-to-oecd-ministers-of-health-from-the-high-level-reflection-group-on-the-future-of-health-statistics_24bc04e2-en.html [Accessed 26 Jun 2025].
12 de Boer D, van den Berg M, Ballester M, *et al*. Assessing the outcomes and experiences of care from the perspective of people living with chronic conditions, to support countries in developing people-centred policies and practices: study protocol of the International Survey of People Living with Chronic Conditions (PaRIS survey). *BMJ Open* 2022;12:e061424.
13 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* 2025;34:529–36.
14 Bloemeke-Cammin J, Groene O, Ballester M, *et al*. International cross-cultural development and field testing of the primary care practice questionnaire for the PaRIS survey (PaRIS-PCPQ). *BMC Prim Care* 2024;25:168.
15 Jung A, Challoumas D, Pagels L, *et al*. Guidelines for the development and validation of patient-reported outcome measures: a scoping review. *BMJ Evid Based Med* 2024;29:363–73.
16 Reeve BB, Wyrwich KW, Wu AW, *et al*. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. *Qual Life Res* 2013;22:1889–905.
17 Porter I, Rijken M, Groene O, *et al*. The International Survey of People Living with Chronic Conditions (PaRIS survey): development of the conceptual framework. *Qual Life Res* 2021;30:S1–177.
18 Valderas JM, Ferrer M, Mendívil J, *et al*. Development of EMPRO: a tool for the standardized assessment of patient-reported outcome measures. *Value Health* 2008;11:700–8.
19 Snape D, Kirkham J, Preston J, *et al*. Exploring areas of consensus and conflict around values underpinning public involvement in health and social care research: a modified Delphi study. *BMJ Open* 2014;4:e004217.
20 University of Michigan. Survey Research Center. Guidelines for best practice in cross-cultural surveys: full guidelines. Survey Research Center, Institute for Social Research, University of Michigan; 2011.
21 Fitzgerald R, Widdop S, Gray M, *et al*. Identifying sources of error in cross-national questionnaires: Application of an error source typology to cognitive interview data. 2011.
22 OECD. OECD-saudi arabia co-operation. Available: https://www.oecd.org/en/about/programmes/saudi-arabia-oecd-co-operation.html [Accessed 21 Jun 2025].
23 Schäfer WLA, Boerma WGW, Kringos DS, *et al*. QUALICOPC, a multi-country study evaluating quality, costs and equity in primary care. *BMC Fam Pract* 2011;12:115.
24 PROMIS. PROMIS list of adult measures. 2023.
25 Topp CW, Østergaard SD, Søndergaard S, *et al*. The WHO-5 Well-Being Index: a systematic review of the literature. *Psychother Psychosom* 2015;84:167–76.
26 Heun R, Burkart M, Maier W, *et al*. Internal and external validity of the WHO Well-Being Scale in the elderly general population. *Acta Psychiatr Scand* 1999;99:171–8.
27 Maibach EW, Weber D, Massett H, *et al*. Understanding consumers' health information preferences: development and validation of a brief screening instrument. *J Health Commun* 2006;11:717–36.

28 Lloyd H, Fosh B, Whalley B, *et al*. Validation of the person-centred coordinated care experience questionnaire (P3CEQ). *Int J Qual Health Care* 2019;31:506–12.
29 OECD. PaRIS patient questionnaire (paris-pq) oecd paris patient questionnaire (paris-pq). 2024.
30 Commonwealth Fund. International health policy survey of older adults.
31 Osborn R, Squires D, Doty MM, *et al*. In New Survey Of Eleven Countries, US Adults Still Struggle With Access To And Affordability Of Health Care. *Health Aff (Millwood)* 2016;35:2327–36.
32 Kruk ME, Lewis TP. Introduction to The Lancet Global Health's Series on the People's Voice Survey on Health System Performance. *Lancet Glob Health* 2024;12:e14–5.
33 Lewis TP, Kassa M, Kapoor NR, *et al*. User-reported quality of care: findings from the first round of the People's Voice Survey in 14 countries. *Lancet Glob Health* 2024;12:e112–22.
34 Lewis TP, Kapoor NR, Aryal A, *et al*. Measuring people's views on health system performance: Design and development of the People's Voice Survey. *PLoS Med* 2023;20:e1004294.
35 Groenewegen P, Spreeuwenberg P, Timans R, *et al*. Data analysis plan of the OECD PaRIS survey: leveraging a multi-level approach to analyse data collected from people living with chronic conditions and their primary care practices in 20 countries. *BMC Res Notes* 2024;17:157.