Patient-reported indicators for assessing health system performance

MEASURING WHAT MATTERS: THE PATIENT-REPORTED INDICATOR SURVEYS

August 2019

OECD
BETTER POLICIES FOR BETTER LIVES

2019 Status Report

Measuring What Matters: the Patient-Reported Indicator Surveys

Patient-reported indicators for assessing health system performance

Acknowledgements

The work was enabled by the financial and substantive assistance of the European Union (grant number 2017 5302) and funding by the OECD Member States.

The authors would like to acknowledge and thank the delegates of the OECD Working Party for PaRIS, the Health Care Quality and Outcomes Working Party, the Health Committee and the PaRIS Taskforce for their thoughtful feedback and comments. The authors would like to acknowledge the members of the condition- and procedure-specific workgroups, who contributed to this work.

This report was prepared by the Health Division of the OECD Directorate for Employment, Labour and Social Affairs.

Table of Contents

Abbreviations

ADL	Activities of Daily Living
CAT	Computer-adaptive testing
HCQO	(Working Party on) Health Care Quality and Outcomes
HLRG	High-Level Reflection Group on Health Statistics
HOOS	Hip dysfunction and Osteoarthritis Outcome Score
IPAQ	International Physical Activity Questionnaire
KOOS	Knee dysfunction and Osteoarthritis Outcome Score
OHS	Oxford Hip Score
PaRIS	Patient-Reported Indicator Surveys
PREM	Patient-Reported Experience Measure
PRIM	Patient-Reported Incidence Measure
PROM	Patient-Reported Outcome Measure
PROMIS	Patient-Reported Outcomes Measurement Information System
QALY	Quality-Adjusted Life year
QOL	Quality of Life
SF-36	SF-36: the Short Form instrument to measure patient health consisting of 36 items
VAS	Visual Analogue Scale

Executive summary

PaRIS helps health systems to become more people-centred

The general objective of the PaRIS initiative is to develop, pilot and implement new patient-reported indicators of health system performance, specifically patient-reported experience measures (PREMs) and patient-reported outcome measures (PROMs). PaRIS helps health systems to become more people-centred by providing systematic, internationally standardized information on what matters most to patients.

The Patient-Reported Indicator Surveys (PaRIS) initiative builds international capacity to measure and compare patient-reported indicators, using indicators that enable comparisons across countries. It also aims to encourage patient-reported indicators to evolve in a common direction internationally, to enable shared learning, development and research.

The OECD Health Committee launched the PaRIS initiative in January 2017. In that month, OECD Health Ministers met in Paris to discuss the next generation of health reforms. These discussions revealed clear political momentum to pay greater attention to what matters to patients. The resulting Ministerial Statement calls on health systems to become more people-centred by developing international benchmarks of health system performance based on data reported by patients themselves.

PaRIS helps policy makers in two ways:

PaRIS comprises two streams of work:

- In areas where patient-reported indicators for specific conditions already exist, the first work stream supports countries to accelerate the adoption and reporting of validated, standardised, internationally comparable patient-reported indicators.
- To address the need to understand the outcomes and experiences of people with one or more chronic conditions, the second work stream is to develop a new international survey. This focuses on adults with one or more chronic conditions who are receiving primary/ambulatory care services.

Condition-specific work: Agreement on performance indicators and pilot data collection of standardised, internationally comparable data

The PaRIS initiative has convened groups of international experts to agree appropriate health system performance indicators to be reported by patients with specific conditions or procedures. Three international working groups have started working on sets of indicators in the areas of hip- and knee replacements, breast cancer surgery and mental health. Using feedback from the working groups, the OECD has piloted data collection of condition-specific performance indicators.

The working group for hip and knee replacements has created a first international database with indicators. This database includes pre- and post-scores on quality of life, measured with the EQ-5D and scores on condition-specific instruments such as:

- the Hip dysfunction and Osteoarthritis Outcome Score (HOOS);
- Oxford hip scores (for hips);
- Knee dysfunction and Osteoarthritis Outcome Score (KOOS);

- and Oxford knee scores (for knees).

Efforts on breast cancer have culminated a pilot data collection involving ten clinical sites from seven countries using postoperative breast satisfaction scale of the breast conserving therapy (lumpectomy) and breast reconstruction modules of the Breast Q tool, an internationally validated instrument used to measure breast surgery outcomes reported by patients.

The working group for mental health has been focusing on both patient reported outcome and patient reported experience measures for persons with serious mental illness. So far, the group has been working through a series of modified Delphi processes to build consensus around the development of international standards for the collection of patient-reported mental health indicators. The Delphi rounds have also focused on the target population for the measure, the time points for PROM and PREM collection, as well as the domains of experience and outcome that should be the priority focus for the international standard. These discussions have revealed key conceptual questions, such as whether to include or exclude priority population groups such as children and young people, or when during the treatment period to collect patient-views, as well as tensions between the areas which are most important, and what is practically feasible.

Implementation plan for the PaRIS International Survey of Patients with Chronic Conditions: a framework with indicator domains and a sampling design developed with an international expert Taskforce

Whereas the first work-stream follows mainly a bottom-up approach, building on indicators that are already measured, for the second work stream a new survey is being developed from the ground up. The survey will target a population of people with one or more chronic conditions who live in private homes and whose conditions are being managed in primary care or other ambulatory care settings. This survey will be multi-staged, so that variation in outcomes could be analysed across countries as well as within countries on different levels (geographical area, provider, patient). The survey implementation plan as well as the content of the survey (indicator domains) has been developed together with the international PaRIS Taskforce. The Taskforce includes representatives from countries interested in participating in the survey as well as selected experts in the domains of PROMs, PREMs, population health survey content, health services survey content, health survey methods and survey operations from academic, clinical, policy and statistical communities.

PROMs that will be measured in the survey relate to physical, mental and social health status. Examples of indicator domains for physical health status are pain, fatigue, physical functioning, sleep and body functions. Examples of indicator domains for mental health are anxiety and symptoms of depression. Indicators domains for social health are social activities, social roles and responsibilities, limitations in social activities, satisfaction with participation in social roles and satisfaction with participation in discretionary activities.

PREMs will be based on ongoing OECD work on patient experience over the past decade. These PREM's cover experiences of care accessibility; quality of the communication between care providers; shared decision-making with care providers; care continuity and coordination; comprehensiveness of care; patient safety; and patient empowerment.

Next steps

The work described in this report forms an important basis for the further development and implementation of PaRIS. For the condition-specific work, this means that further

standardisation and harmonising reporting of existing data, either by the use of similar instruments and procedures or by creating crosswalks between different instruments. For the International Survey of Chronic Diseases, next OECD will develop a concrete study protocol, operation manuals and a questionnaire. This will be done based on review of the scientific literature and by consulting academic experts, patients, providers, and other relevant stakeholders.

- ***Study Protocol:*** In order to make data internationally comparable, it is important that international guidelines for the survey, for instance with regard to the eligible population, conditions to be included, etc. are adhered to. However, there are many differences across health systems and many practical aspects will need to be tailored to the national or regional situation. Such differences concern for instance: the most effective way to identify and recruit providers, the relevant stake holders to involve and alignment with existing initiatives. Some countries already have national surveys in place that have many similarities with the PaRIS design. In such cases, creating synergy between the two projects may be more efficient than starting a second survey from scratch. PaRIS will work with experts who are knowledgeable of the national context and with strong networks in the national community to make a national study protocol. This protocol describes in detail the steps to be taken for a successful implementation and all relevant national stake holders and how to involve them.
- ***Operations Manual:*** Based on the national protocol, PaRIS will develop its operations manual, a practical manual that describes all procedures in detail and that includes materials to be used in the survey, such as standard letters to inform, recruit, or remind providers and patients, guidelines for a safe exchange of data, etc. This work will include all activities related to the development of the instruments (questions and scales) and the questionnaire to be used in the field trial.
- ***Questionnaire:*** The questionnaire will be suitable for online-surveying. For patients that are unable to fill out a questionnaire online, the questionnaire can also be used for telephone or face to face surveying. When filling out the questionnaire online, most patients will be able to finalize it within 20-30 minutes. PaRIS carry out cognitive testing among small groups of patients.

Additional next steps for the PaRIS initiative will finalise the methodological development required to establish the agreed set of Patient-Reported Incidence Measures (PRIMs) for international data collection. The OECD will develop indicator specifications and guidelines in preparation for pilot testing through national and international surveys is use in selected member countries. This work on PRIMs may be potentially integrated into other PaRIS work streams, as well as the work of the OECD health division more broadly.

Introduction

In the past years, the urgency of putting people more in the centre of health systems has gained momentum across the OECD. The time has come for a new generation of health system indicators that reflect what health systems really deliver to the people they serve.

In 2015, the OECD convened a High-Level Reflection Group on Health Statistics to advise how the Organisation could enhance its quantitative assessment of health system performance. The Group recommended that OECD develop patient-reported indicators on outcomes and experiences of care, with a focus on international comparison, by developing indicators both in areas where the OECD already collects other data; and filling in gaps in such indicators in primary care, where currently no indicators are available.

On the first area, the Group recommended that work begin with validated metrics already in use and seek to accelerate international adoption and/or harmonisation across countries. Patient priorities and perspectives should be taken into account when developing and prioritising indicators. Where valid patient-reported metrics do not yet exist for priority diseases, sectors or services, new indicators and patient surveys should be developed.

In January 2017, OECD Health Ministers met in Paris to discuss the next generation of health reforms. These discussions revealed clear political momentum to pay greater attention to what matters to patients. The resulting Ministerial Statement calls on health systems to become more people-centred by inter alia developing international benchmarks of health system performance based on data reported by patients themselves.

Responding to the recommendations of the High-Level Reflection Group on Health Statistics and the Ministerial mandate, the OECD launched the Patient Reported Indicator Surveys (PaRIS) initiative. PaRIS aims to build international capacity to measure and compare care outcomes and experiences as reported by patients, using indicators that enable comparisons across countries. It also aims to encourage patient-reported indicators to evolve in a common direction internationally, to enable shared learning, development and research. PaRIS comprises two work streams:

- ***Work stream 1:*** In areas where patient-reported indicators such as PROMs and PREMs already exist, the first work stream supports countries to accelerate the adoption and reporting of validated, standardised, internationally comparable patient-reported indicators.
- ***Work stream 2:*** To address the need to understand the outcomes and experiences of people with one or more chronic conditions, the second work stream is to develop a new international survey. This focuses on adults with one or more chronic conditions who are receiving primary/ambulatory care services.

The first work stream builds on existing country, and sub-national level data collection, using this data to develop internationally comparable patient-reported indicators. The second work stream involves a new international survey that is being developed from the ground up.

This report describes early results from the work that has been done within the framework of PaRIS in the first two years. Since PaRIS in a comprehensive, large-scale international initiative, with much innovative and pioneering work, it will take several years and the results described in this report should be seen as preliminary results. The results as reported

are based on in-depth research and many intensive interactions with a range of world-leading experts. They provide a solid basis to build further on this important initiative.

More specifically, this report describes the following main elements of PaRIS:

- ***Chapter 1:*** This chapter describes how measures on outcomes of care and experiences with care as reported by patients play a key role in the transformation towards people-centred health systems. Such measures can be used on different levels and with different objectives, for instance for benchmarking, monitoring, quality improvement, peer-learning and to foster the dialogue between patient and provider. The chapter also gives a 'taxonomy' of outcomes and experiences and provides some definitions. This chapter is common to the work of both work streams.
- ***Chapter 2:*** This chapter describes the background and rationale of the PaRIS initiative and why and when it was launched. It describes the objectives, main work streams, of PaRIS. The chapter also describes seven key principles that will guide all work within the PaRIS initiative. This chapter is common to the work of both work streams.
- ***Chapter 3:*** This chapter describes a number of measures recommended to be used to measure outcomes of care for people who undergo joint replacements, knee replacement, breast cancer surgery and people with mental health conditions. Three international expert working groups developed these recommendations using a bottom-up approach. The chapter will also discuss the possibility to make crosswalks between instruments in order to enable comparisons of different registries that use different tools to measure similar domains. The chapter also contains results of the pilot data collection. This chapter relates to work stream 1.
- ***Chapter 4:*** This chapter describes the key indicator domains to be used to measure outcomes of care and experiences with care of patients with (multiple) chronic conditions who are largely managed in primary care or other ambulatory care settings. For each indicator domain, a number of existing indicators are defined and several pros and cons of different instruments are discussed. This chapter relates to work stream 2.
- ***Chapter 5:*** This chapter describes the sampling methodology for the survey of patients with chronic conditions. More specifically, it will discuss the definitions of the population of eligible patients and providers, how a multi-stages sampling will be implemented and how providers and patients will be recruited in such a way that satisfactory response rates will be realised. This chapter relates to work stream 2.
- ***Chapter 6:*** This chapter describes common conclusions for both work streams.

Chapter 1. Measuring what matters for people-centred health systems

This chapter describes how measures on outcomes of care and experiences with care as reported by patients play a key role in the transformation towards people-centred health systems. Such measures can be used on different levels and with different objectives, for instance for benchmarking, monitoring, quality improvement, peer-learning and to foster the dialogue between patient and provider. The chapter also gives a 'taxonomy' of outcomes and experiences and provides some definitions.

1.1. We need to know how health care and health policy affect the lives of people

The primary objective of any health system, service or organisation is to maximise the health of individuals and the populations they serve, to do so in an equitable way within budgetary parameters.

Good health is not just important in its own right. It also promotes personal, social and economic well-being. Healthy people create healthy communities and contribute towards a well-functioning, prosperous and more productive society. For example, good health can enhance a person's lifetime earnings by up to 25% (OECD, 2018[1]) (OECD, 2017[2]).

Yet very few health systems assess their impacts on health and well-being from the perspective of the people they serve. While the concept of health-related quality of life (QoL) has existed for almost three decades, it is not measured or reported systematically. Performance metrics in health tend to focus principally on inputs and outputs. Outcomes such as life expectancy are important, but they are silent on a range of other things valued by patients, including pain, function and QoL as well as the experience of care itself. This means that the picture of health care and health system performance is missing an essential component.

The patient perspective is increasingly relevant in overcoming the demographic, epidemiological and economic challenges faced by all health systems. The rise of chronic conditions as the main source of disease burden, coupled with better technologies to manage them and prolong life, heightens the need for a more people-centred approach to both policy and practice.

1.2. A people-centred health system needs to measure what matters to patients

People's assessment of their health, and the outcomes of care, go beyond whether they survive a disease or medical intervention. A range of inter-related physical and mental health domains including pain, mobility, fatigue, anxiety and symptoms of depression all contribute to person's health-related QoL. Patients also value their care experience: Was my autonomy respected? Was I able to participate in decisions about my care? Did the care pathway - its organisational aspects - place an undue cost or burden on me?

It makes sense to capture this knowledge in a way that is systematic and useful for decision-making. Yet the health sector has been remiss at measuring the effects of its activities on outcomes and experiences as reported by patients. Forward thinking provider organisations, in disease registries and in some health systems have been collecting this

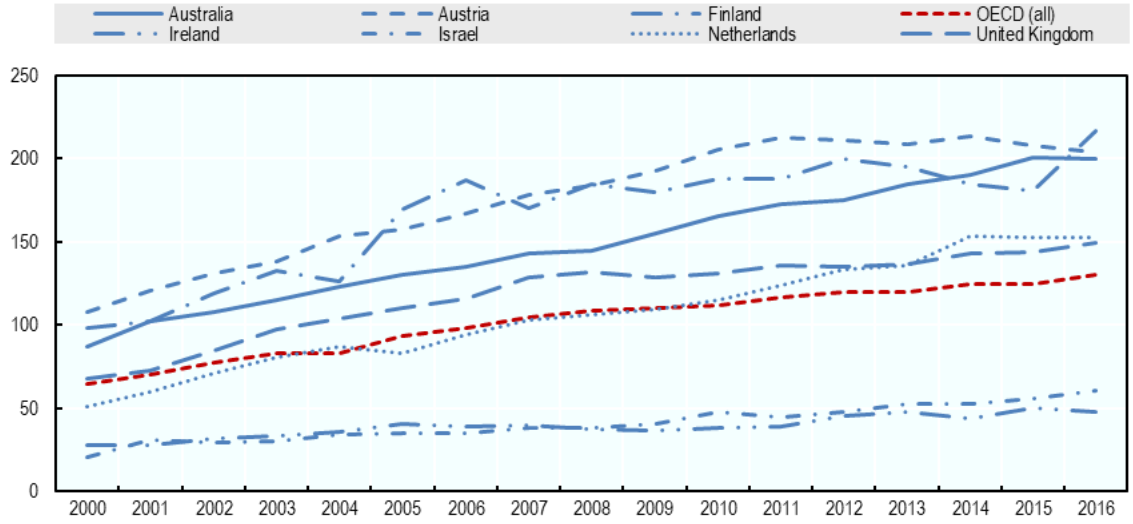
information for some conditions or procedures, however, coherent and systematic patient reporting across the entire range of health system activities is not yet in place.

1.2.1. Outputs provide only a partial picture of health system performance

This is in stark contrast to processes and activities, which are routinely collected and reported, yet – in isolation -- reveal quite little about quality and value. For example, the average rate of total knee replacement in OECD countries doubled between 2000 and 2016 (Figure 1.1). Rates also vary up to 5-fold between and within countries (OECD, 2014[3]). Are the increased rates and the variation warranted? Do these operations make a difference to people's lives, or are some of them performed unnecessarily? What is the effect of waiting times for, and patient's age at surgery? Are some patients better off choosing other available treatments?

Figure 1.1. Total knee replacement rates have doubled since 2000

Total knee replacement rates per 100,000 population adjusted for population ageing, selected countries and all-OECD average



Note: OECD (all) is the age-adjusted rate of all countries that submit data on this procedure. Countries were chosen based on data availability over this period
Source: https://stats.oecd.org/

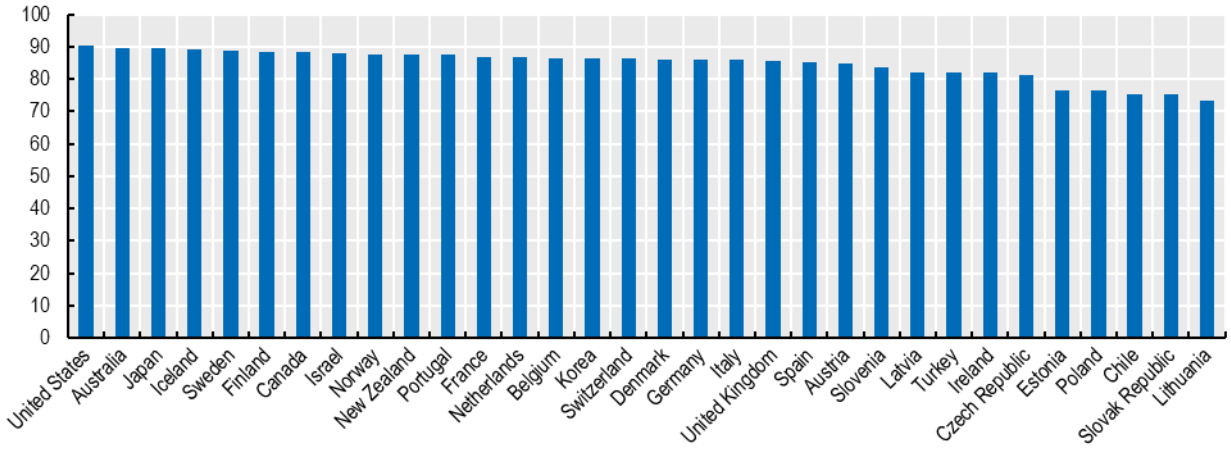
Such questions cannot be answered without knowing care outcomes. Case fatality or hospital re-admission are useful measures but are becoming rare in routine procedures. They are also silent on other outcomes valued by these patients, notably reduction of pain, increase of mobility and function.

1.2.2. We know how medicine treats diseases; but how are people treated?

Traditional outcome measures like survival or mortality will remain useful but cannot capture more subtle yet important effects. For example, people diagnosed with cancer value survival highly. However, therapeutic success entails more (Abahussin et al., 2019[4]). Survival and mortality say little about nausea, pain, sleep quality, body image, sexual function, independence and time spent with loved ones. Also, for many conditions, mortality and survival rates are quite converged (Figure 1.2), with little separating the 'best from the rest'. This hinders continued learning about best therapeutic approaches, techniques and interventions (Hamdy et al., 2016[5]), (Donovan et al., 2016[6]).

Figure 1.2. Cancer survival has converged

Breast cancer age-standardised five-year survival 2014



Source: https://stats.oecd.org/

That medicine has become quite successful at treating disease should be celebrated. However, continual improvement must include assessment of impact treatments have on people's lives. This makes outcomes valued by patients a key indicator of success. Most men diagnosed with prostate cancer are now also very likely to survive this condition. But they also highly value preserving erectile function and avoiding incontinence (Nag et al., 2018[7]) -- outcomes of significant interest to patients, providers as well as policymakers.

1.2.3. A good care experience contributes to better clinical outcomes, it is also an end in itself.

In addition to outcomes, *how* people are treated also matters. This includes being treated with respect and compassion and being supported, listened to and involved in decision-making. It also means that care is better integrated across teams who communicate well with each other and with the patient.

A positive care experience is a strong signal of quality care and is instrumental in outcomes achieved, especially for those who manage multiple chronic conditions (Stein et al., 2015[8]) (Trzeciak et al., 2016[9]) (Luxford, Safran and Delbanco, 2011[10]). In mental health, for example, a positive care experience influences the relationship with the care team, manifesting in better communication, therapeutic continuity, adherence and health outcomes (Wong et al., 2019[11]). But it is also an important end in itself. All patients expect and deserve to be treated with respect. In some sectors, such as palliative care, aspects of the care experience such as compassion and dignity are among the most important components of care.

Yet despite considerable progress in some specific cases, the care experience is not captured systematically. This needs to change, given the growing importance of this dimension of service delivery.

1.2.4. Patient-reporting supports shared decision making

In the clinical setting, measuring patient-reported metrics helps to focus the health care interaction on the needs of the individual. The discussion moves from 'what's the matter

with you?' to 'what matters *to* you?' – a critical first step in shared decision making, a core principle of people-centred care. Aggregate patient-reported outcomes can inform care decisions and help choose the right therapeutic option where various interventions (including watchful waiting') are available (Veroff, Marr and Wennberg, 2013[12]). People see what the most likely outcomes of an intervention may be and can decide accordingly.

Regular reporting by patients throughout their care journey adds structure and rigour to assessment, decision-making and action. Care can be better tailored to individual needs, and enables a rapid and accurate response to clinical deterioration. For example, reporting of symptoms by patients during chemotherapy has been found to significantly prolong survival and improve quality of life (Basch, 2017[13]) (Basch et al., 2017[14]).

Knowledge derived from patient-reported data can be used to develop decision aids and update clinical practice guidelines. It also informs providers on how their work affects patient health and well-being. Patient-reported outcomes, for example, provide a way to measure clinical progress more objectively. It can complement other metrics to provide a fuller assessment of their performance. If implemented well, benchmarking and even public reporting can be a powerful driver of quality improvement (Greenhalgh et al., 2018[15]).

Data generated by patients can also contribute towards assessing the performance of medical products, combination therapies, care pathways, health services and the health system as a whole. Combined with other data this can furnish researchers, regulators, health technology agencies, payers, academics and policy makers with the knowledge to make more informed decisions to maximise health system performance, and meet the expectations of patients, citizens and communities (Calvert, O'Connor and Basch, 2019[16]).

1.3. Patient-reported indicators are robust and reliable

The ability to elicit information from individuals on their health status, quality of life and care experience is now decades old. Many instruments and surveys have undergone rigorous psychometric testing and statistical validation, with results published in the peer-reviewed literature. The field is mature and evidence supports that these instruments reliably measure what is intended (Black, 2013[17]). Box 1.1 outlines the different types of Patient-reported outcome and experience measures (PROMs and PREMs) as well as some of the technical aspects of how these are collected, interpreted and used.

Box 1.1. Measuring patient-reported outcomes and experiences of care

Typically, instruments to elicit information from patients on self-reported health status, outcomes and experiences of care comprise questionnaires of varying length and format. These are administered in a range of ways (verbally, electronically or on paper). The two main categories of patient-reported outcome measures (PROMs) are condition-specific PROMs and health-related Quality of Life instruments – commonly termed 'generic' PROMs.

Condition-specific PROMs

These instruments are designed specifically for a condition (e.g. osteoarthritis) or a procedure (e.g. joint replacement). These PROMs are tailored to the symptoms of a specific condition, or those that a specific procedure tries to address. As such their advantage is sensitivity and specificity.

Their key limitation is a lack of generalisability – that is, their results cannot be directly compared with other conditions or procedures.

Health-related QoL instruments ('generic' PROMs)

'Generic' PROMs instruments attempt to capture a broader range of physical and psychosocial domains that are considered important determinants of health-related QoL. Their advantage is that they can be compared across different conditions, procedures and interventions. They are often used in health technology assessment (HTA).

Patient-reported experience measures (PREMs)

Patient experience is also measured using surveys or questionnaires. These can be administered in various ways. A number of approaches and questions have been developed. Questions can be tailored to a certain setting (e.g. primary, hospital, long-term care) or assess a specific aspect of care (e.g. continuity, autonomy, information provision). PREMs are now sophisticated and often rooted to objectives events, having moved well beyond the more subjective patient 'satisfaction' surveys of the past. They elicit scaled data across a range of dimensions including accessibility, communication, continuity and confidence. These data are now used to inform assessment and international comparisons of health systems **(Doty et al., 2017[18])**.

Collecting and using patient-reported data

A range of factors influence the outcomes of care as reported by patients, including behaviour, adherence, age and comorbidities. But readmission and mortality are subject to the same confounding variables. Like any outcome data that are used for benchmarking, confounders for patient-reported indicators should usually be adjusted in order to enable meaningful comparisons **(Nuttall, Parkin and Devlin, 2015[19])**. All data, whether patient-reported or not, have limitations and should be interpreted with the necessary caution.

In the end, no single data source can provide information for a complete assessment of how a complex, adaptive health system performs. Patient-reported data need to be interpreted in the context of other metrics on health system activity and performance. They are not meant to supplant but to complement existing data that are collected to avoid tunnel vision and generate a more complete picture of performance for all involved from patients and providers to regulators and policy makers.

In order for patient-reported indicators to fulfil their promise in service provision, research and policy, standardisation of methods for data collection, analysis and reporting are essential. This relies heavily on international collaboration (Calvert, O'Connor and Basch, 2019[16]).

Chapter 2. The Patient Reported Indicator Surveys: objectives, rationale and key principles

This chapter describes the background and rationale of the PaRIS initiative and why and when it was launched. It describes the objectives, main work streams, of PaRIS. The chapter also describes seven key principles that will guide all work within the PaRIS initiative.

2.1. Background and rationale

2.1.1. Policy makers today know little about what healthcare systems deliver to patients

Health systems are in need for better information about the value and outcomes they produce. There is little information available about the impact of health care services upon the people served, beyond re-admissions to hospital, complications and deaths. There is a strong need to assess health care outcomes from the perspective of the people served. This is key to learning how well health services deliver their ultimate objective: supporting people in regaining and sustaining their health and well-being.

Given the global trend of increased expenditure on health care as a share of national income, it is surprising that systematic, empirical measurement of the outcomes and experiences of care from the patient's perspective is still the exception in most healthcare systems. This gap in knowledge limits the ability for evidence-based policy making and the ability to maximise benefits of health care at acceptable costs. It is difficult to improve what is not being measured. The PaRIS initiative addresses this gap in knowledge.

Patient-reported outcome measures are in use for some conditions, such as hip and knee surgery, but different measures and methods in different countries make international comparisons difficult. Moreover, the most rapidly growing group of health care users are seldom asked about their outcomes and experiences. These are people with one or more chronic conditions who are managed in primary care or other ambulatory care settings.

2.1.2. Health ministers asked the OECD to lead an effort to fill this gap in knowledge

In January 2017, Health Ministers of OECD Member countries asked the OECD Health Committee to lead an effort to develop and analyse cross-country comparative measures of patients' own experiences of medical care and health care outcomes. This mandate draws from the recommendations of a High-Level Reflection Group on Health Statistics (HLRG), convened by the Health Committee in 2015.

The final report of the HLRG addressed the need for more information on patient-reported experiences and outcomes of care to better monitor health system performance and drive continuous improvement. The report stressed that several patient-reported experience measures (PREMs) already exist internationally, including those collected by OECD and reported in the Health at a Glance publication, but coverage and comparability remain limited. The challenge is even more apparent for patient-reported outcome measures (PROMs). Whilst a number of PROMs exist (particularly for hospital-based procedures) in a selection of countries, the existence of multiple concurrent initiatives risks missing the

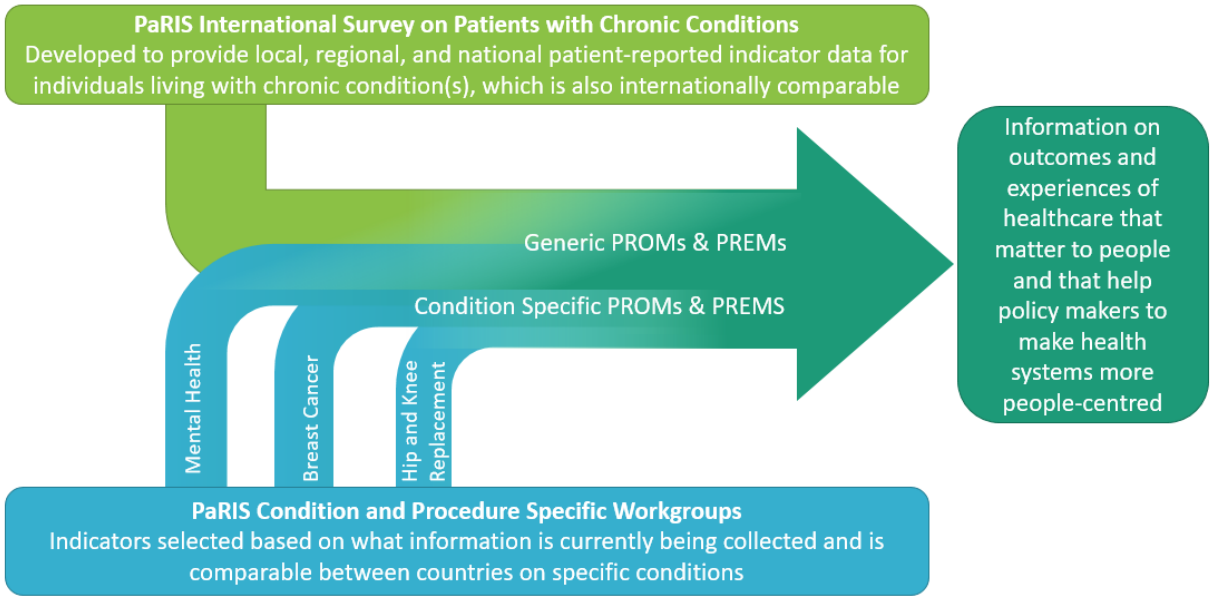
opportunity for international comparison. Of greater concern, PROMs are noticeably absent from the fastest growing area of health care need (and spending): routine care for people with chronic conditions. Measuring experiences and outcomes reported by these patients is an important element to assess whether care is delivering what people need.

2.1.3. PaRIS contains two streams of work:

PaRIS complements the data already collected by the OECD on health system performance, with information on patient-reported outcomes and experiences. PaRIS will also fill the information gap in neglected areas that pose emerging health system challenges, particularly around the effective management of people with chronic conditions.

- Work stream 1: In areas where patient-reported indicators such as PROMs and PREMs already exist, the first work stream supports countries to accelerate the adoption and reporting of validated, standardised, internationally comparable patient-reported indicators. Three international working groups have started in early 2018 to discuss instruments (questions, or scales containing multiple questions) measuring patient-reported outcomes, definitions and data collection strategies in three areas: breast cancer, hip and knee replacements and mental health.
- Work stream 2: To address the need to understand the outcomes and experiences of people with one or more chronic conditions, the second work stream is to develop a new international survey. The survey focuses on of adults with one or more chronic conditions who are receiving primary/ambulatory care services.

Figure 2.1. PaRIS' two complementary work streams together work to provide useful policy information on health system performance



Source: Authors.

2.2. Objectives of the PaRIS initiative

In this section, the objectives of the PaRIS initiative in general and of the two working streams will be outlined.

2.2.1. Objective of PaRIS

The general objective of the PaRIS initiative is to develop, pilot and implement new patient-reported indicators of health system performance, specifically patient-reported experience measures (PREMs) and patient-reported outcome measures (PROMs). PaRIS helps health systems to become more people-centred by providing systematic, internationally standardized information on what matters most to patients.

2.2.2. Objective of the PaRIS work on specific conditions

There is a gap in the information required to enable patients, providers, policymakers and others make more informed decisions. The PaRIS work on specific conditions aims to accelerate the uptake of existing patient-reported indicators for specific conditions, in part by harmonizing existing national level data collection efforts. This PaRIS work is twofold: the first part of this work convenes international experts to come to agree on appropriate health system performance indicators to be reported by patients with specific conditions such as breast cancer or those who received hip or knee surgery. Secondly, the initiative aims to pilot the collection of data for identified condition-specific performance indicators. Member states have ongoing opportunities to review and comment on the results.

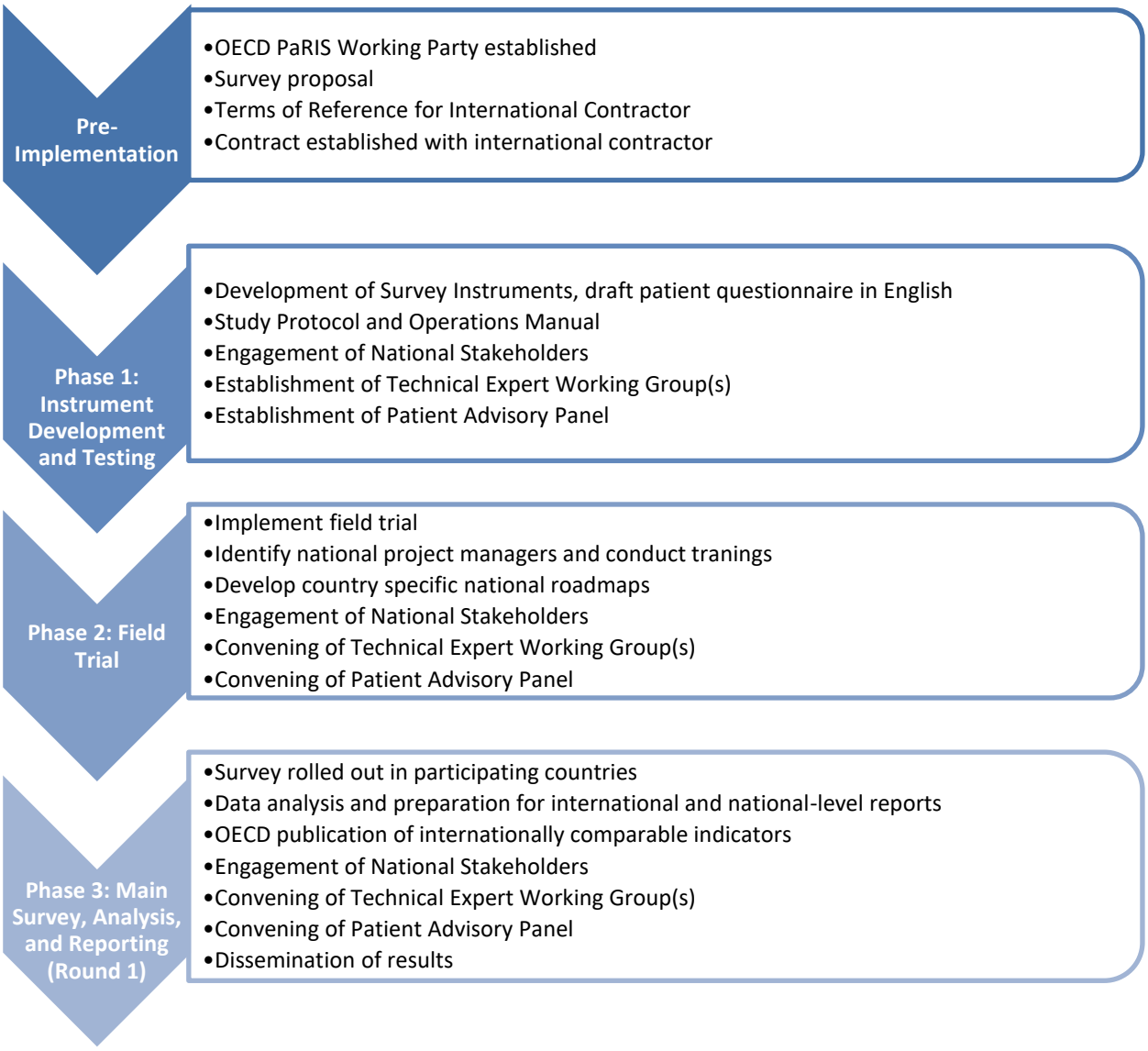
More specifically, each of the condition or procedure specific Working Groups does the following:

- Recommend suitable instrument(s) for international collection of patient-reported outcomes for these procedures and, oversee assessment of appropriate crosswalks between instruments to support comparable reporting.
- Develop indicator definitions, specifications and standards for comparable reporting (inclusions, exclusions, collection time points, and risk adjustment protocols) and a minimum data set for collection.
- Develop standards and best practice guidelines for international data collection. This will include guidance on sampling requirements, collection methods (e.g. electronic) and privacy and security of data.
- Advise on requirements to ensure comparability of results between languages and cultures.
- Advise on international benchmarking and reporting requirements, and on the publication of data.
- Advise on implementation support such as training manuals, protocols, and stakeholder engagement, especially patients and clinicians.
- Share information on high-level resource requirements for PROMs collection, and potential approaches to improve efficiencies.
- Share national and international experience in this domain.

2.2.3. Objectives and Implementation Plan for the PaRIS survey of patients with chronic conditions

The main objective of the PaRIS survey is to report internationally comparable health care outcomes and experiences of adults with chronic conditions who are treated in primary/ambulatory care through indicators reported by patients themselves and that can be repeatedly measured over time. This means that the indicators selected should highlight variation in key outcomes within and across countries to support national health care system improvement. Since very little has been measured for this group of patients so far, a new international survey is being developed from the ground up (see Figure 2.2).

Figure 2.2. PaRIS International Survey on Chronic Disease Process

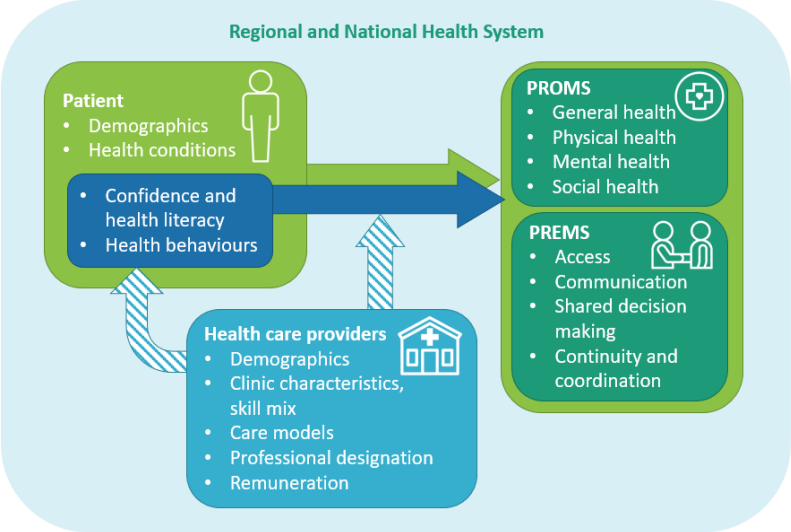


The PaRIS survey offers an opportunity to gather the evidence necessary to transform health systems to be centred on the needs of the people they serve. Comparing the performance of health systems will inform policy makers and help them understand to what

extent their policies are on track to deliver more people-centred health systems. This will also shed light for people with chronic conditions, helping them to understand how the outcomes and experiences of care in their own country compare with those in other countries; and it will help open a dialogue with service providers about how to further improve the performance of health services and health systems to become more people centred.

The PaRIS conceptual framework links the three main data domains of the survey (see Figure 2.3). The first is data information about the patient provided by the patient. This includes information about their demographics and health conditions, as well as information about levels of patient engagement, health literacy, and healthy behaviours.

Figure 2.3. PaRIS Survey Conceptual Framework



Note: Clarification: variation in PROMs and PREMs reported by patient will vary based on their background characteristics and health conditions; these will be included as case mix variables or for analysing specific subgroups. Patients' confidence and health literacy will also affect proms and PREMs, these factors are modifiable and could be influenced by health care. The relation between patient characteristics and the PROMs /PREMs they report could be influenced by health care providers. Patients, providers and the relations all exists in and are influence by the regional and national health system.
Source: Authors.

The second group of data is generated from primary/ambulatory care providers. These providers are the first and main contact point between community-dwelling adults with chronic conditions and the health care system. Primary/ambulatory care services play a pivotal role in the continuous treatment of people with chronic conditions; support people to self-manage their care; and advise and assist people on their health care pathway that may involve many other parts of the health care system. They are, therefore, the most appropriate part of the health care system through which to reach adults with chronic conditions who are receiving health care in their community and to understand their experiences throughout the health care system. In some countries, routine care for chronic conditions is typically provided in primary care facilities, for instance by, or under the supervision of, general practitioners or family physicians, whereas in other countries other care settings, such as outpatient clinics of hospitals play an important role. Health care

providers influence both a patient's engagement and behaviours, and have an influence on patient-reported outcomes.

The final group of data is the PROMs and PREMs themselves. This patient-reported data is provided by the patient, and relates to their demographic and health conditions, as well as patient's health literacy and behaviour. The health system and health providers dually influence these outcomes.

The creation of the PaRIS International Survey of Chronic Diseases builds on pervious OECD work (see Box 2.1)

Box 2.1. Building on previous OECD work on patient experience measures in population based surveys

As part of the work of the Working Party on Health Care Quality and Outcomes, the OECD Secretariat has made progress in measuring and reporting patient experiences (PREMs) in an internationally comparable manner since 2006. The Secretariat developed a set of questions which could be embedded in national population-based health surveys. This includes questions about access to health care and patient experiences with primary/ambulatory care. These questions have become part of HCQO's regular data collection. The number of countries reporting these data for international comparisons has been steadily increasing over time.

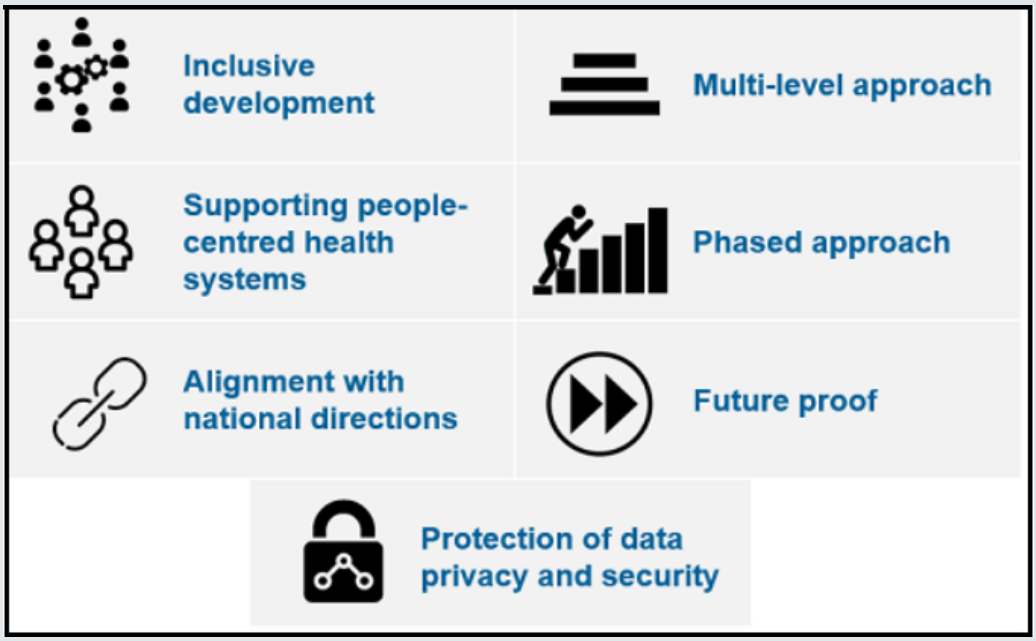
Most of these data have been reported in previous editions of Health at a Glance (2013, 2015, and 2017). This set of questions has been modelled after the Commonwealth Fund International Health Policy surveys. Several countries that collect data from the Commonwealth Fund surveys have also started to collect these data through population-based national health surveys in order to improve the quality and response rates, and to undertake in-depth analyses. The questions have also been included in national population-based health surveys in an increasing number of countries that are not covered by the Commonwealth Fund surveys.

There is a strong value in further expanding the scope of current PREMs measures collected via population surveys, for example to include other dimensions of experiences important for health care users and by adding a limited number of generic PROMs. However, survey design and sample size constraints restrict the possibility to report the experiences of patients with chronic conditions or to relate them to characteristics of health care systems.

2.3. Guiding principles of PaRIS

A number of key principles will guide the development of the PaRIS survey, and the PaRIS initiative more broadly. These guiding principles reflect the Ministerial Statement that provided the mandate for this work, discussions that took place with many stakeholders, and the unique features of these types of indicators (see Box 2.2).

Box 2.2. Guiding principles of PaRIS



1. Inclusive development: together with all stake holders and countries
2. Supporting more people-centred health systems: information that is actionable and that helps to identify policy actions to improve care
3. Alignment with national directions and initiatives: creating synergy with initiatives already going on in countries
4. Multi-level approach: combining information on the level of patients, health care organisations and health systems to get the full picture.
5. Phased approach: development, field trial, implementation. Countries will we involved in key decisions in all phases.
6. Future-proof: use of state-of-the art, innovative methods for data collection and data sharing that are safe, privacy-respectful and user-friendly.
7. Protection of data privacy and security

These principles are as follows:

2.3.1. Inclusive Development

Instruments and indicators that are being developed should be relevant and valuable for policy makers, patients and health care providers. These key stakeholders will be involved in all stages of the project. Inclusive development also means that as many countries as possible are involved in the development of instruments to assure that cultural differences

and differences between systems are taken into account. All participating countries take part in key decisions.

2.3.2. Supporting people-centred health systems

PaRIS will produce information that enables policy makers and other stakeholders to understand variation in health outcomes and health care experiences and to identify policy actions to improve care. PaRIS supports ongoing efforts in many countries to strengthen the involvement and empowerment of patients in their own health care. The collection of data is only a means to this goal.

PaRIS also promotes the embedding of patient-reported indicators in the care process to enable providers to use the information for quality improvement, and enable patients and providers to use the information to enhance shared decision-making and communication about the outcomes of care during health care encounters. For the patients and providers involved, participation should be as easy as possible in terms of time and energy required and constitutes a positive experience.

2.3.3. Alignment with national directions

In several countries, patient reported indicators are already being collected. PaRIS builds on this experience and experts involved in such national initiatives are advising the Secretariat. To the extent possible, implementation of PaRIS will be aligned with national initiatives that are already underway. This will, for example, be done by creating consistency and comparability in the use of generic PROMs and alignment with initiatives to embed PROMs and PREMs in the national/regional data infrastructure.

2.3.4. Multilevel approach

Factors that influence patient experiences and outcomes can be identified at different levels: the level of patients, the level of health care professionals, the level of health care organisations and the level of health systems. Therefore, it is desirable that patient-level data can be linked to these higher levels and that variation on higher levels could be analysed. Depending on the system, regional levels may also be included.

2.3.5. Phased approach

Following the successful approach used by other international data collections, such as the OECD Programme for International Student Assessment (PISA) and the OECD Programme for International Assessment of Adult Competencies (PIAAC), PaRIS will progress through several phases and countries are asked to commit phase by phase. Examples of these phases are development – field trial – implementation of main survey.

2.3.6. Future proof

Data collection takes place preferably electronically, with a user-friendly and safe interface. State-of-the-art solutions are being used to implement the survey data collection and feedback.

2.3.7. Protection of data privacy and security

The survey design and the practices of data processors must fully protect the privacy of survey participants, both patients and health care providers; and must adhere to all applicable legislative requirements for data privacy and security protection and follow the principles within the OECD Recommendation on Health Data Governance (OECD, 2015[20]). The survey design and the practices of data processors must fully protect the privacy of survey participants, both patients and health care providers; and must adhere to all applicable legislative requirements for data privacy and security protection and follow the principles within the OECD Recommendation on Health Data Governance (OECD, 2015[20]).

Chapter 3. Patient reported measures for the assessment of care for specific conditions: hip replacements, knee replacements, breast cancer surgery and mental health

This chapter describes a number of measures recommended to be used to measure outcomes of care for people who undergo joint replacements, knee replacement, breast cancer surgery and people with mental health conditions. Three international expert working groups developed these recommendations using a bottom-up approach. The chapter will also discuss the possibility to make crosswalks between instruments in order to enable comparisons of different registries that use different tools to measure similar domains. The chapter also contains results of the pilot data collection.

3.1. Joint replacement rates are rising but are patients reporting improvement?

Each year, over 2.2 million people undergo an elective hip or knee replacement in OECD countries. Knee replacement rates have doubled since the year 2000 (Figure 1.1), while hip replacements have increased by 30%. Inter- and intra-country variation in rates can be as high as 5-fold (OECD, 2014[3]).

Patients typically undergo these procedures to manage symptoms of osteoarthritis such as pain and loss of mobility and function, which impact on health-related QoL. Both procedures are invasive and, like all surgery, involve a degree of risk. They require a long period of rehabilitation. They are also expensive. In Australia, for example, they account for over 2% of total health expenditure.[1]

Given that alternative ways of managing hip and knee pain exist (physical therapy, exercise and medication) patients should be able to base their decision to proceed with surgery on the expected outcomes including pain, mobility and capacity to perform daily activities following a period of recovery. Payers should expect that the procedures represent value compared to the alternatives.

The orthopaedic community has been one of the most active in encouraging the collection of patient-reported data. Nevertheless, national-level reporting is the exception. Most patient-reported data collections are part of regional and local programmes, or voluntary registries covering a subset of a country's providers and hospitals. In addition, a range of instruments measuring dimensions such as pain, function and QoL are in use around the world. The instrument is typically completed by the patient pre-surgery and then at a specified time point after the operation (usually 6 or 12 months). The numerical difference between the two scores is the key value of interest.

The OECD has been working with a range of stakeholders and experts, including patients and clinicians, to collect PROM data internationally. Ten programmes across eight countries contributed to a recent pilot data collection. These included national initiatives (England, Netherlands, Sweden), regional (Canada - Alberta and Manitoba, Switzerland - Geneva), sub-national registries (Australia – ACORN) and single hospitals (Finland – Coxa hospital,2 Italy – Galeazzi Institute). Various PROM instruments are used among the contributing programmes, and the post-operative data collection time points vary 6 months.

Adult patients with a diagnosis of osteoarthritis3 who underwent a unilateral, primary elective total replacement procedure were included in the data collection. Three most recent

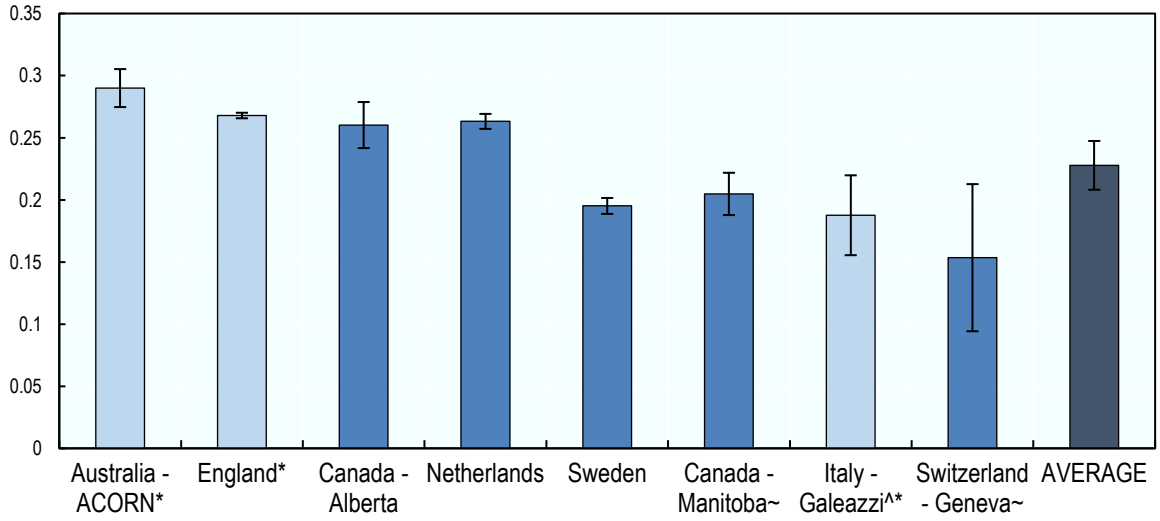
years of data were collected and aggregated to provide one result per participating programme.

3.1.1. Hip replacement patients reported improvement on average

Hip replacement results derived from the generic instruments (EQ-5D-3L, EQ-5D-5L and SF12) are presented on a common metric -- the EQ-5D-3L index with a United States-derived valuation (Shaw JW, 2005[21]) (see Box 3.1). On this scale, the maximum score is 1.0 (denoting optimal health-related QoL) while a negative score suggests health-related QoL rated as worse than death.

Figure 3.3 presents the difference between the mean pre- and mean post-operative scores[4] -- i.e. the mean change in QoL -- adjusted for patients' age, sex and pre-operative score. Results suggest that the average patient in each programme reported improvement in their health-related QoL following a hip replacement. The average mean adjusted change across the programmes was +0.23, which equates to approximately 21% improvement on this index.[5,6]

Figure 3.1. Hip replacement: adjusted mean change adjusted mean change between pre- and post-operative EQ-5D-3L scores (US valuation), with 95% confidence intervals, 2013-16 or nearest years



Note: ^ results converted from SF-12v1 instrument; ˜converted from SF-12v2 instrument; *6-month post-op collection (all others are 12 months); *Source*: PaRIS Hip/Knee Replacement Pilot Data Collection

Box 3.1. The EQ-5D index and data standardisation

The EQ-5D health-related QoL instrument comprises questions covering five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. The patient rates each from 1-3 (on the 3L version) or 1-5 (on the 5L version) with 1 being best and 3 or 5 worst. The output is a five-digit 'health state' – e.g. 11111 (perfect health), 33333 or 55555 (worst possible state for 3L and 5L respectively) and a range of permutations in

between. The instrument also asks the respondent is also asked to place a marker on a 0 - 100 visual analogue scale (VAS) to indicate their current state of health.

The health states are converted to a single index by referring to so-called valuations specific to a population or country. These valuations have been determined by asking a sample of that population about how they would rate a particular health state against being in perfect health (1.0) and death (0) using a method called time trade-off (TTO). The resulting function is called a valuation or value set.

Currently, over a dozen national valuations exist for the 3L version, but fewer have been completed for 5L – a newer version. The functions can differ considerably between countries (Zhuo et al., 2018[22]). Some remain above zero but many decline into negative values at the worst possible health states. This means respondents rated these states as worse than death, and were willing to trade off time in good health to avoid that health state.

The EQ-5D was designed to generate quality-adjusted life years (QALYs) - a measure that combines morbidity and mortality and is often used assess the effectiveness of medical interventions. For example, living in a health state of 0.8 on the index for 10 years equates to 8 QALYs.

The EQ-5D-3L index (US valuation) as the common scale

The EQ-5D-3L index was chosen as the common metric because (a) the majority of countries use this instrument; (b) algorithms exist to convert – or map – scores from other generic instruments to the EQ-5D-3L. Score conversions were conducted using patient-level data.

'Native' EQ-5D-3L health state valuations (see above) exist for most participation programmes. A single valuation, rather than a mix of respective native value sets, is preferred because it goes some way to mitigate cultural, demographic, socio-economic and other confounders of self-reported health status (Devlin, 2019[23]). It de facto presents results consistent with their underlying health state, and removes the additional variability created by a country's unique valuation of these states.

The choice of the US valuation was pragmatic. It was the only 'end point' of the available algorithms to generate EQ-5D-3L scores from the other instruments used by the contributing programmes (van Hout et al., 2012[24]) (Sullivan and Ghushchyan, 2006[25]) (Le, 2013[26]).

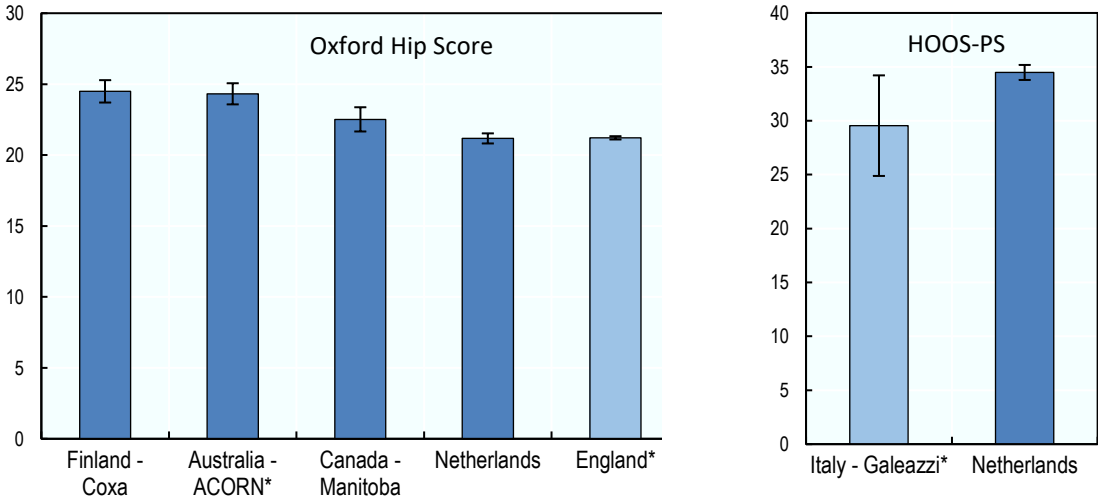
Standardising results to enhance comparability

To enhance comparability and mitigate the effect of demographic and other variables, results shown (derived from both generic and condition-specific tools) were adjusted for age, sex as well as the reported pre-operative PROM score, based on the pooled data of the contributing programmes. Three age categories and two pre-operative score categories were used. Differences between crude and adjusted results were small in the majority of cases. Results were not adjusted for co-morbidity or socio-economic status due to the lack of consistent data.

The adjusted changes between pre-and post-operative scores derived from condition-specific instruments (Oxford Hip Score, HOOS-PS)[7] are presented in Figure 3.2. They must be displayed on separate axes because algorithms to convert scores are not available at

present. Results suggest, on average, improvement in all programmes. The Oxford scale ranges from 0 to 48, the HOOS-PS from 0 to 100. In both cases a higher value represents a more desirable outcome.[8] Results are quite similar across the programmes. The average adjusted mean change (not shown) was +23 on the Oxford scale and +32 on the HOOS-PS scale, which equates to about 48% and 32% improvement respectively[9] (more condition-specific results are provided in Chapter 6).

Figure 3.2. Hip replacement: adjusted mean change between pre- and post-operative Oxford Hip Score and HOOS-PS scores with 95% confidence intervals, 2013-16 or nearest years



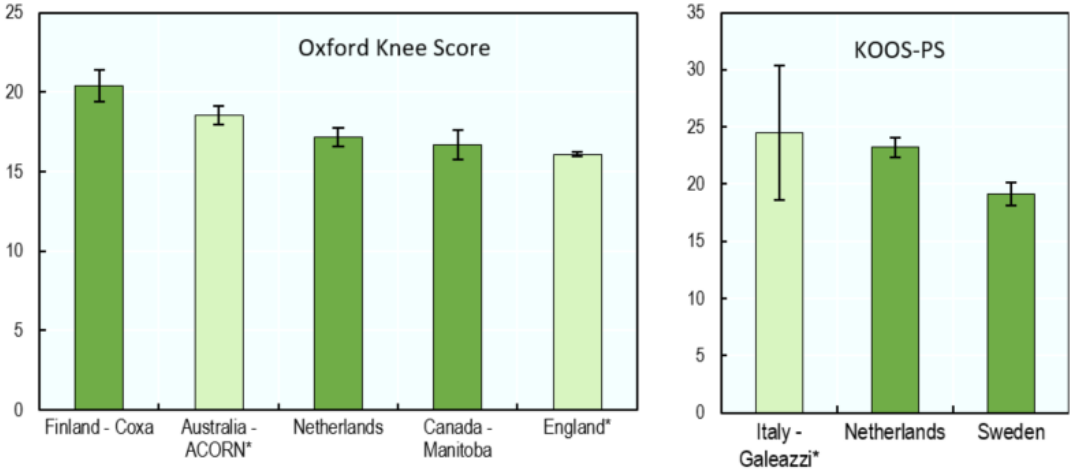
Note: *post-op collection at 6 months (all others at 12 months); Scales: Oxford 0-48; HOOS-PS 0-100

Source: PaRIS Hip/Knee Replacement Pilot Data Collection

3.1.2. *Knee replacement patients reported more modest improvements*

The adjusted changes between pre-and post-operative knee replacement scores derived from condition-specific instruments are presented in Figure 3.3 (the scales are the same as for hip replacement). On average, patients in each programme reported improvement. The average adjusted mean change (not shown) was +17 on the Oxford scale and +22 for KOOS-PS,[10] or 36% and 22% improvement respectively (the corresponding values for hip replacement were 48% and 32%).[11]

Figure 3.3. Adjusted mean change between pre- and post-operative Oxford Knee Score and KOOS-PS scores with 95% confidence intervals, 2013-16 or nearest

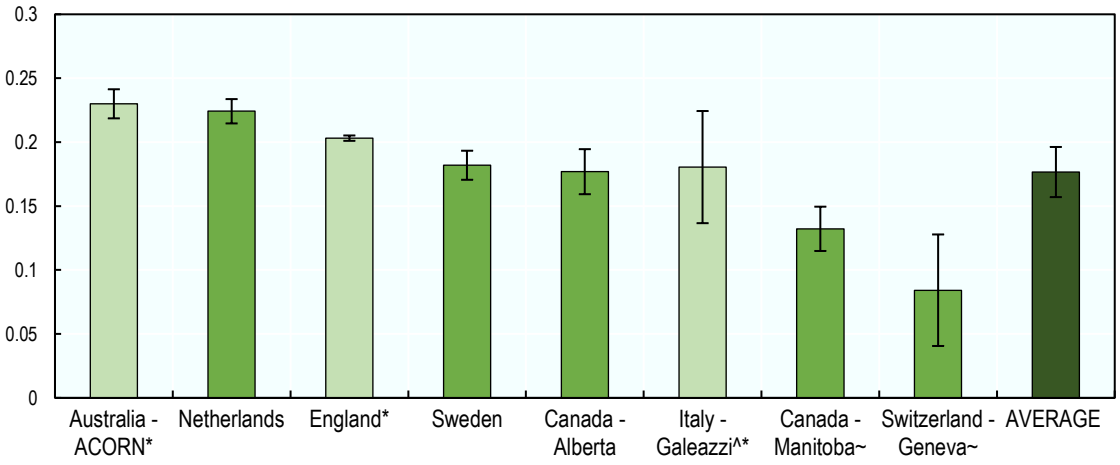


Note: *post-op collection at 6 months (all others at 12 months); Scales: Oxford 0-48; KOOS-PS 0-100

Source: PaRIS Hip/Knee Replacement Pilot Data Collection

Knee replacement results derived from generic instruments are presented using the EQ-5D-3L index with US valuation (see Box 2). Data derived from EQ-5D-5L and SF-12 scales were converted using validated algorithms (van Hout et al., 2012[24]) (Sullivan and Ghushchyan, 2006[25]) (Le, 2013[26]). Figure 3.4 shows the mean changes between pre- and post-op scores, adjusted for age, sex and pre-operative score. Here the average patient in each programme reported improvement, ranging from +0.08 to +0.22. The average adjusted mean change across all programmes was +0.18 (about 16% improvement). The hip replacement equivalent value was +0.23 (21%),[12] a statistically meaningful difference at the 95% level.

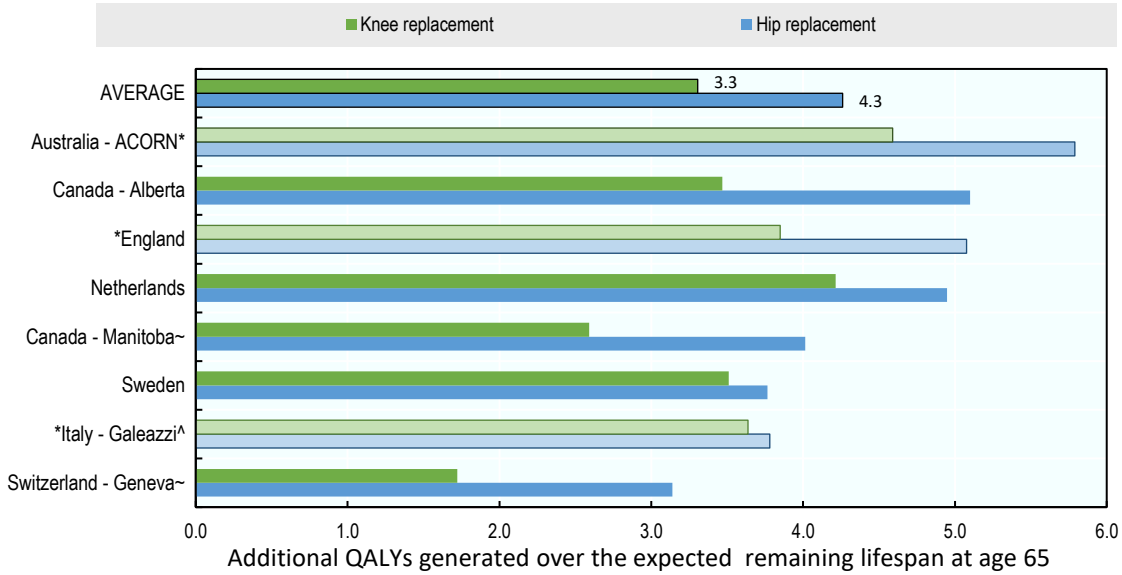
Figure 3.4. Knee replacement: adjusted mean change adjusted mean change between pre- and post-operative EQ-5D-3L scores (US valuation), with 95% confidence intervals, 2013-16 or nearest years



Note: ^ results converted from SF-12v1 instrument; ˜converted from SF-12v2 instrument; *6-month post-op collection (all others are 12 months); *Source*: PaRIS Hip/Knee Replacement Pilot Data Collection

The EQ-5D results suggest that -- all other things being equal – the average patient undergoing a knee replacement at age 65 in the contributing programmes gained an additional (incremental) 3.3 quality adjusted life years (QALYs).[13] In other words, they gained the equivalent of 3.3 years in 'full' health-related QoL over the expected remainder of their life (compared to a 'no intervention' alternative). The corresponding figure for hip replacement is higher 4.3 QALYs (see Figure 3.5).[14] The difference between the procedures is consistent with existing literature (Konopka et al., 2018[27]).

Figure 3.5. Both hip and knee replacements generate additional QALYs for patients



Note: ^ results converted from SF-12v1 instrument; ˜converted from SF-12v2 instrument; *6-month post-op collection (all others are 12 months) *Source*: PaRIS Hip/Knee Replacement Pilot Data Collection

3.1.3. Results should be interpreted with caution

On average, patients undergoing hip or knee replacement procedures in the participating programmes reported, an improvement in their symptoms and health-related QoL. This does not mean that all patients improved. In fact, a small but significant proportion reported no change or a worsening in their symptoms and health-related QoL for both procedures across the participating programmes. While this may still represent a better outcome compared to the counterfactual of no intervention, this is unlikely given the availability of other treatment modalities. Results presented here are in fact silent on how the outcomes of hip and knee replacement surgery compare with other, more conservative, medical treatments for joint pain. This would require expanding the study cohort to patients who choose non-surgical treatments.

Although results were standardised for age, sex and pre-operative score, a number of programme-specific variables limit their comparability. The number of patients differ considerably. Some of the contributing programmes collect post-operative scores at 6 months, others at 12 months. The latter is considered to be the optimal time for post-operative assessment as full recovery is expected 1 year after surgery. Programmes deploy

different modes of collecting data (paper, electronic, telephone) which is known to influence results. The response rates vary between programmes. Despite adjustment for pre-operative score, differences in wait times between countries may also influence results. Finally, results from three programmes were converted from, EQ-5D-5L and SF-12 to the EQ-5D-3L index (US valuation), which may bias the final results.

Results have not been adjusted for case mix and co-morbidities because consistent data were not available across all programs. Moreover, a range of cultural, demographic and socio-economic factors influence self-reported health status and will influence the comparability of results, even when a common index and valuation are used

3.1.4. Better information on breast cancer care outcomes can help patients with difficult treatment choices

Breast cancer is the most prevalent form of cancer in women, with about 2.1 million newly diagnosed cases in 2018 accounting for almost 1 in 4 cancer cases among women (Bray et al., 2018[28]). While an increase in the incidence of breast cancer over the past decade has been observed, mortality has declined in most countries. Early diagnosis as well as improved treatments have contributed, with 5-year net survival rates of 80% now evident in most OECD countries.

A range of medical and surgical treatment approaches exist. The three main surgical pathways are:

- **Breast conserving therapy** or **lumpectomy** involves a surgical operation to remove the cancer while leaving as much of the breast as possible – commonly an option in early-stage cancer. This is the primary surgical choice for breast cancer, with 60%–80% of newly diagnosed cancers amenable to breast conservation at diagnosis or after primary systemic therapy for women in Western Europe (Cardoso et al., 2019[29]).
- **Mastectomy** involves complete removal of the breast surgically and is often undertaken when a woman cannot be treated with a lumpectomy. However, a woman can choose a mastectomy over a lumpectomy for personal reasons and women at very high risk of getting a second cancer sometimes have both breasts removed.
- **Breast reconstruction** may be chosen by women, who have had surgery as part of their breast cancer treatment, to rebuild the shape and look of the breast. The two main types of breast reconstruction are: 1) **implant** reconstruction surgery which involves the insertion of a silicone implant after the removal of the woman's breast tissue; and 2) **autologous** reconstruction surgery, which uses tissue from other parts of the woman's body, such as her stomach, back, thighs, or buttocks to rebuild the breast shape. This form of reconstruction is generally considered to look more natural and behave more like natural breast tissue than breast implants.

The choice of surgical intervention influences the chances of survival as well as subsequent quality of life. Women diagnosed with breast cancer can therefore face difficult decisions when considering treatment options. While factors such as age, general health status and the size and location of primary tumour are important to clinical decision making, the preferences of the patient are central to the choice of treatment strategy (Cardoso et al., 2019[29]). Amidst the overarching objective to stay alive, a key consideration is QoL.

In weighing up treatment options, information about the outcomes of other women who have been in similar circumstances can potentially be of great help in the decision making process and ongoing reflection of progress during and after treatment.

3.1.5. The capacity to collect and use PROMs in breast cancer care is growing

Motivated providers and patients across OECD countries are increasingly measuring patient-reported care outcomes to help inform difficult clinical decisions. The utility of such measurement is increasingly appreciated. For example, in the Netherlands breast cancer has been identified as one of the possible priority areas as part of a current national policy effort to measure patient reported outcomes systematically. Nevertheless, a variety of different PROM tools are used, making comparability of outcomes more difficult, and the scale of uptake is still largely localised and isolated to specific initiatives and clinical champions at specific sites.

In an effort to reflect this emerging priority, the OECD worked with a group of experts (including patients, clinicians, policymakers and industry representatives) and collaborating organisations to understand the current state of the art in breast cancer PROMs and to explore opportunities for international data collections and comparisons.

These efforts have culminated a preliminary international data collection involving 10 clinical sites from 7 countries (Flinders Medical Centre, Adelaide, Australia, Charité – Universitätsmedizin Berlin, Germany, Erasmus Medical Center, Rotterdam, Netherlands, Capio St Göran Breast Unit, Södersjukhuset Bröstcentrum and Karolinska Univ.sjukhuset Bröst Endokrin och Sarkom, Stockholm, Sweden, Universitätspital Basel, Basel, Switzerland, Manchester University Hospitals NHS Foundation Trust, Manchester, UK, Memorial Sloan Kettering Cancer Center, New York, US and Brigham and Women's Hospital, Boston, US) using postoperative breast satisfaction scale of the BCT (lumpectomy) and breast reconstruction modules of the Breast Q tool, an internationally validated instrument used to measure breast surgery outcomes reported by patients (Pusic et al., 2009[30]) (see box 3.2). The data collection involved women aged 15 years and older who received a unilateral lumpectomy or breast reconstruction following a mastectomy during the primary treatment of breast cancer.

Box 3.2. Breast Q Postoperative Breast Satisfaction Scales

The Breast Q suite of tools is one of the more widely used amidst the range of instruments currently is use internationally to measure patient reported outcomes from breast cancer surgery (Tevis et al., 2018[31]).

The breast satisfaction scales of the Breast Q tools measure body image in terms of a woman's satisfaction with her breasts and asks questions regarding how comfortably bras fit and how satisfied a woman is with her breast area both clothed and unclothed. Postoperative items ask about breast appearance (e.g., size, symmetry, softness), clothing issues (e.g., how bras fit; being able to wear fitted clothes) and location and appearance of scars. There are separate modules for lumpectomies, mastectomies and reconstructions, with each module consisting of multiple separate scales covering such issues as psychosocial wellbeing, sexual wellbeing, physical wellbeing, satisfaction with breasts and satisfaction with care. There are also implant-specific items, including the amount of rippling that can be seen or felt.

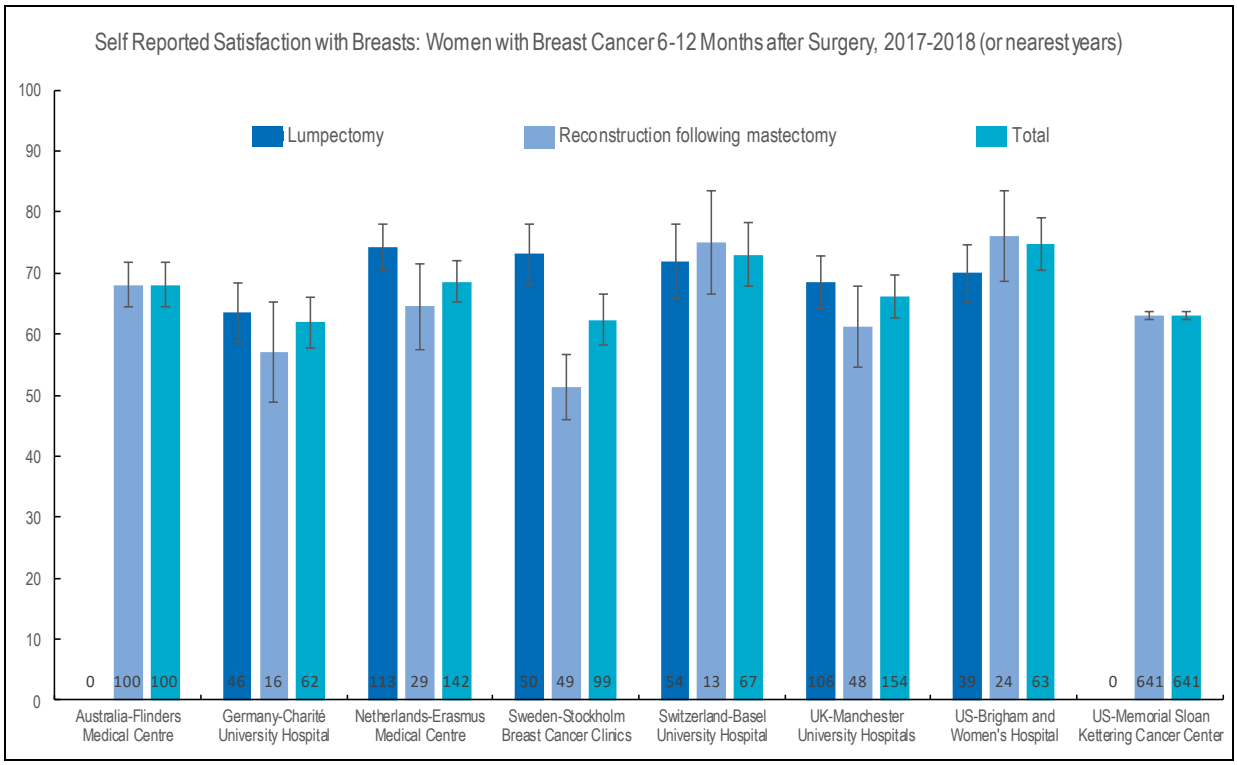
The scores from each scale of the Breast Conserving Therapy and Reconstruction scales, along with the other Breast Q scales can be transformed to an Equivalent Rasch Transformed Score of 1-100 to allow direct comparison between scales.

See http://qportfolio.org/breast-q/breast-cancer/ for more details.

3.1.6. Results suggest satisfaction with breast surgery outcomes are related to the type of procedure

The crude (unadjusted) outcomes scores at 6-12 months following lumpectomy, breast reconstruction, and the aggregate of the two are provided in Figure 3.6. Results are not intended to be representative of the outcomes of breast cancer patients across each country but do show the capacity for metrics of this kind to be reported internationally. While crude data from sites that reported scores for lumpectomy and reconstruction suggest that women in most sites may have higher breast satisfaction outcomes after a lumpectomy, aligning with conventional wisdom in this area, for some sites it appears that women may have higher reconstruction scores. Further work and more extensive data collection are needed to validate this observation and consider the generalizability of the data outcomes, but these early observations may provide some basis for further sharing and learning of outcomes across sites.

Figure 3.6. Crude PROM scores for breast cancer point to variations in surgical outcomes



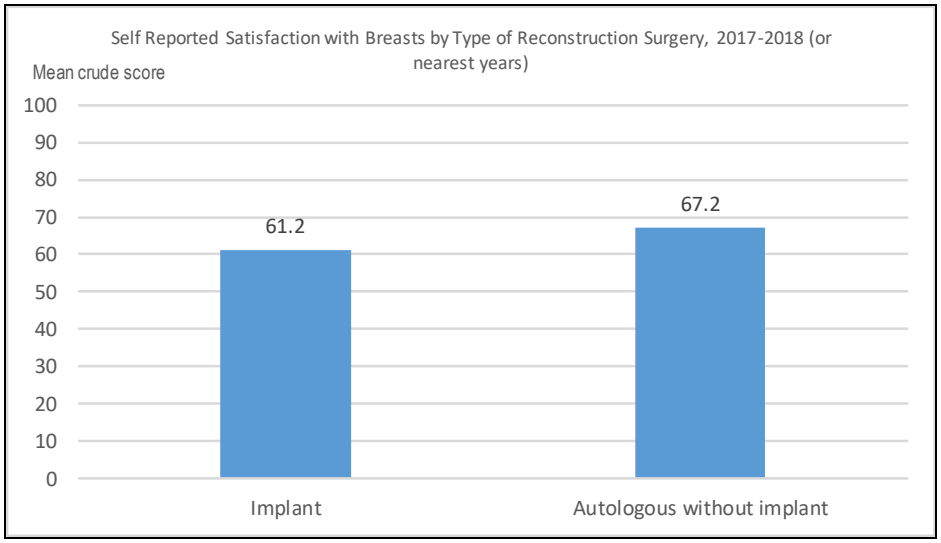
Note: Measurement extended beyond 12 months after surgery for sites in both Sweden and Switzerland. The data labels at the base of the histogram refer to the sample size at each site
Source: PaRIS Breast Cancer PROMs Pilot Data Collection, 2019.

A number of personal factors can influence a woman's postoperative satisfaction with the outcomes of her breast cancer surgery, including age, smoking, obesity, tumour burden and overall satisfaction with breasts and physical health before surgery. For example, smoking and obesity can impair tissue healing and have a negative impact on implant reconstruction results, including aesthetic outcomes (Kern et al., 2015[32]). These factors are largely outside of the health services direct influence and their impact should ideally be taken into account when comparing the quality of care across sites. Data were collected from participating sites on key patient variables, including age, smoking and obesity but limitations on sample size and incomplete capacity for reporting by all sites prevented risk-adjusting results for the time being

3.1.7. Women report slightly more satisfaction with tissue based reconstruction compared to silicon implants

Consolidated crude scores from the participating sites indicate that women are about 10% (6 percentage points) more satisfied with their breasts after autologous reconstruction surgery than women after implant reconstruction (see figure 3.7) This result aligns with existing evidence (Matros et al., 2015[33]) and can be an important consideration where choice of surgical intervention is possible.

Figure 3.7. Crude patient reported outcomes for implants and autologous reconstructions



Note: Derived from consolidated data from all 10 participating sites
Source: PaRIS Breast Cancer PROMs Pilot Data Collection, 2019

It follows that the variation in breast satisfaction scores presented in Figure 3.6 may be influenced, amidst other factors, by the proportion of women undergoing autologous reconstruction surgery. Table 3.1 presents the sample size of women and the proportion undergoing autologous reconstruction reported by each site. The proportion ranges from 100% of women receiving autologous reconstructions (Dutch and Swiss sites) to 0% in the Swedish site, where all women would have received implant reconstructions. While no clear relationship between the proportion of women undergoing autologous reconstruction and the overall crude outcomes scores (see Table 3.1) is apparent, further consideration of the factors contributing to the observed wide variation across sites may be warranted given the conventional wisdom regarding care outcomes, For example, the role of each site within

the broader service arrangements for women with breast cancer and the representativeness of the data.

Table 3.1. Total breast reconstructions and the proportion of autologous reconstructions by site

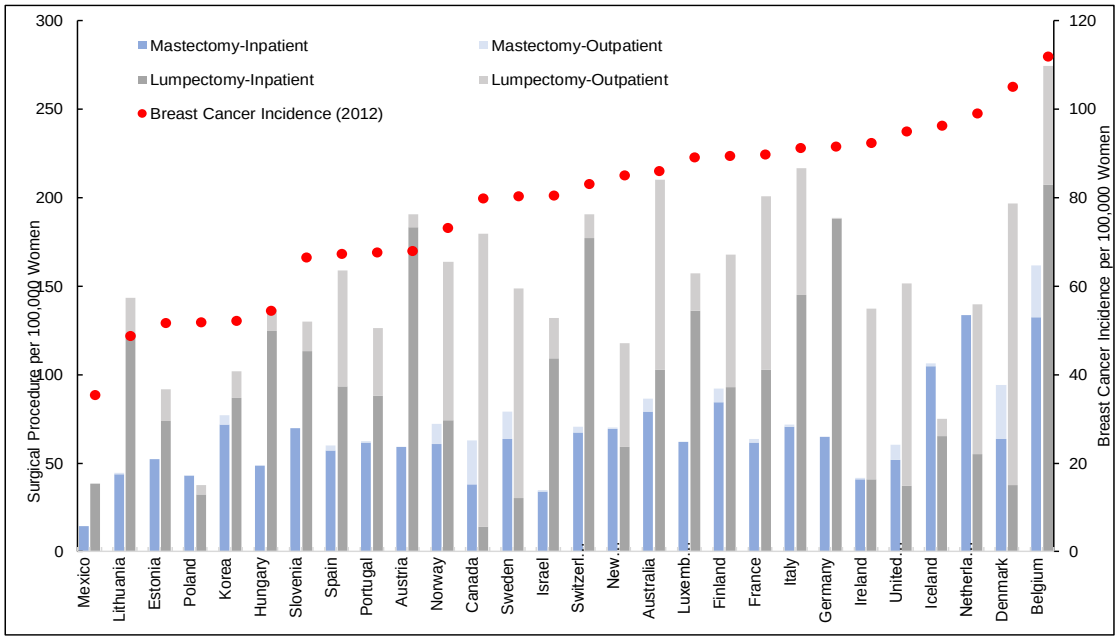
	Total breast reconstructions	Autologous reconstructions without implant
	No. of Women	% of total reconstructions
Australia-Flinders Medical Centre	100	57%
Germany-Charité University Hospital	16	19%
Netherlands-Erasmus Medical Centre	29	100%
Sweden-Stockholm Breast Cancer Clinics	49	0%
Switzerland-Basel University Hospital	13	100%
UK-Manchester University Hospitals	48	25%
US-Brigham and Women's Hospital	24	38%
US-Memorial Sloan Kettering Cancer Center	641	24%

Source: PaRIS Breast Cancer PROMs Pilot Data Collection, 2019

3.1.8. Recent use of PROMs indicates that tissue based reconstruction may be cost-effective

Significant variation in treatment pathways and practices persists for patients with breast cancer, including the use of different surgical approaches, even in the face of established clinical practice guidelines (Cardoso et al., 2019[29]; OECD, 2013[34]). Figure 3.8 presents the rates, setting and mix of lumpectomy and mastectomy surgery across OECD countries, demonstrating that quite different treatment patterns are being pursued, even across countries showing a very similar level of cancer incidence. Data need of course to be interpreted cautiously as patients' cancer stages, comorbidity and pre-operative patient performance status may also vary.

Figure 3.8. Breast cancer incidence, surgery type and setting per 100 000 women, 2017



Source: OECD Health Data 2019, www.oecd.org/health/healthdata

Treatment choices made by patients in consultation with their clinical teams have not only consequences for survival and QoL, but also financial implications. For example, after a mastectomy a woman faces the choice of whether to have breast reconstruction or not and if she proceeds with breast reconstructive surgery, what type of reconstruction she should have. While the outcomes in terms of survival of having a breast reconstruction or not after a mastectomy are generally comparable (Platt et al., 2015[35]), the choice of reconstruction can lead to different outcomes (e.g. quality of life, satisfaction with breasts) for a woman, as well as different health system cost.

While autologous reconstructions appear to result in better patient outcomes than implant surgery, they tend to be more expensive. This raises questions over value for money (Scurci et al., 2017[36]). A recent study in the United States compared the Breast Q scores of patients who had implants and those who underwent autologous reconstructions. This study calculated the average additional cost for obtaining 1 year of perfect breast-related health for a unilateral autologous reconstruction at just under USD12k, compared with implant reconstruction, with lower additional costs for younger patients and earlier stage breast cancer (Matros et al., 2015[33]).

Although, society's value for a year of perfect breast-related quality of life is unknown, a threshold of USD 50k to USD100k for a year in perfect overall health has commonly been used to identify cost-effective interventions and considered as acceptable for adoption of new technologies or techniques in developed countries such those in the OECD (Cameron, Ubels and Norström, 2018[37]). On this basis, there are indications that autologous reconstructions may warrant consideration as a cost effective option to implant reconstructions.

Routine collection of data on outcomes that matter for breast cancer patients are useful not only for direct patient care but also for system improvement through better understanding

of the impact of different care pathways. They complement traditional measures such as survival, mortality, complications and readmissions. Bringing measures of what matter to patients into the equation creates potential to evaluate alternative modes of treatment both in terms of outcome and value for patients, policy makers and third party payers (Cardoso et al., 2019[29]). Measuring the care experience of people with mental health conditions helps improve the quality of care

3.1.9. Existing mental health measures say little about experiences and outcomes of care

Mental health is a vital component of individual well-being as well as social and economic participation. However, many OECD countries consider that their mental health care is inadequate. Between one in six and one in five people experience a mental health problem in any given year, and an estimated one in two people experience a mental health problem in their lifetime (Institute for Health Metrics and Evaluation, 2019[38]). The most common mental health problems are anxiety disorder (5.1 % of the population), followed by depressive disorders (4.5 %), and drug and alcohol use disorders (2.9 %) (ibid.).

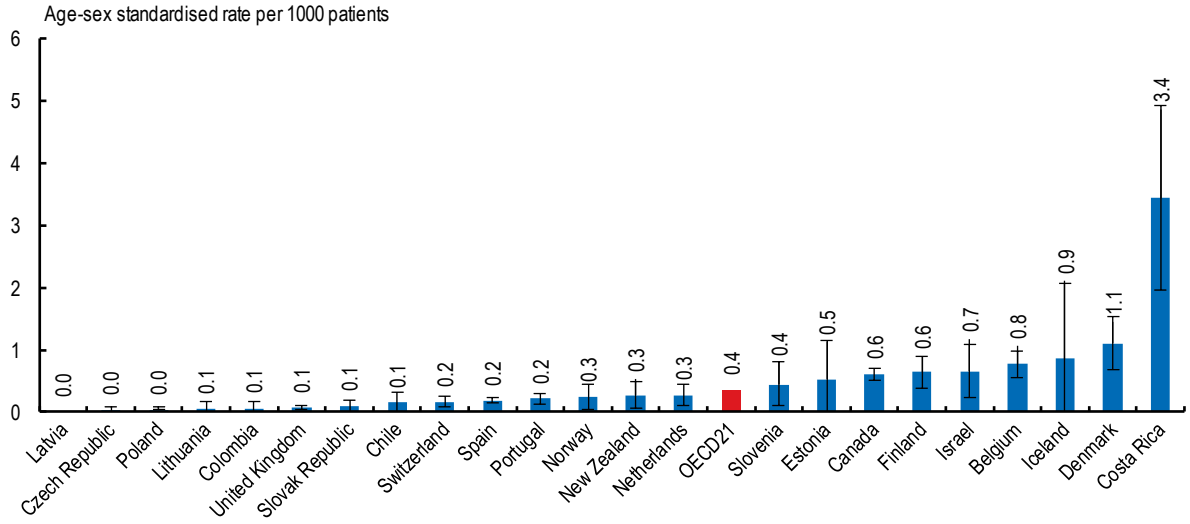
The economic and social costs of mental ill-health are also significant. Direct spending on mental health services was estimated to account for around 13% of total health spending – or 1.3% of GDP – across EU countries in 2015 (OECD/EU, 2018[39]). But larger costs are also borne outside of the health system. Lower employment rates and productivity of people with mental health issues incur economic impact equivalent to 1.6% of GDP in EU countries; with greater spending on social security programmes, such as disability benefits or paid sick leave, accounting for a further 1.2% of GDP (OECD/EU, 2018[40]).

Comparable cost estimates have been established in OECD countries beyond the EU. In Australia, for example, the total costs of mental ill-health amount to 4% of GDP, 45% of which are indirect costs (Australian Government - National Mental Health Commission, 2016[41]), Similar figures are reported in Canada and Japan (Sado et al., 2013[42]; Sado et al., 2013[43]; Mental Health Commission of Canada, 2012[44]).

The impact of mental health problems on individuals' lives, and on societies and economies, can be addressed through more effective policies and interventions to prevent and manage them. However, understanding of the impact that mental health care makes on service users' lives is still weak; there is a pressing need to measure the effects and impact of prevention and treatment approaches more consistently and methodically.

Traditional measures say little about the lasting impact that mental health care has on the patient. For example, inpatient suicide is a critical safety measure which indicates when something has gone terribly wrong with the quality of a patient's care (see Figure 3.9), and one of the limited measures of care quality that can currently be reported internationally. However, inpatient suicide is thankfully very rare, which means for the vast majority of psychiatric patients we do not have a meaningful insight into their experience or outcomes of care.

Figure 3.9. Inpatient suicide amongst patients with a psychiatric disorder, 2015-2017



Note: Three year average except for New Zealand.
Source: OECD Health Statistics 2019.

Patient-reported measures are a critical tool for improving policy and practice in mental health care. An example of how patient-reported measures (in this case PREMs) can shed light on potential problems with mental health care is provided Figure 3.10, which report survey data on the care experience of people who report having mental health conditions compared to those who do not.

Survey data can provide some early insights into the care experience of care for people with a mental health problem

Patient-reporting can shed light on how successful services and health systems are in ensuring people receive quality care for their mental health problems as well as other conditions and comorbidities. The OECD is working with patients, clinicians and policymakers to develop mental health PREM and PROM data collection that enable international comparisons. In the meantime, new analyses of existing data can provide interesting insights.

Although the overall results of the Commonwealth Fund's 2016 International Health Policy Survey of Adults are well known, the comparison of patients with a mental health problem to the rest of the surveyed population has not been conducted and brings new insights on the experience and outcomes of patients affected by mental health problems. While the survey data has several limitations (see Box 3.3), some patterns in the experiences of persons who have (or had) a mental health problem emerge.

Box 3.3. The Commonwealth Fund International Health Policy Survey of Adults

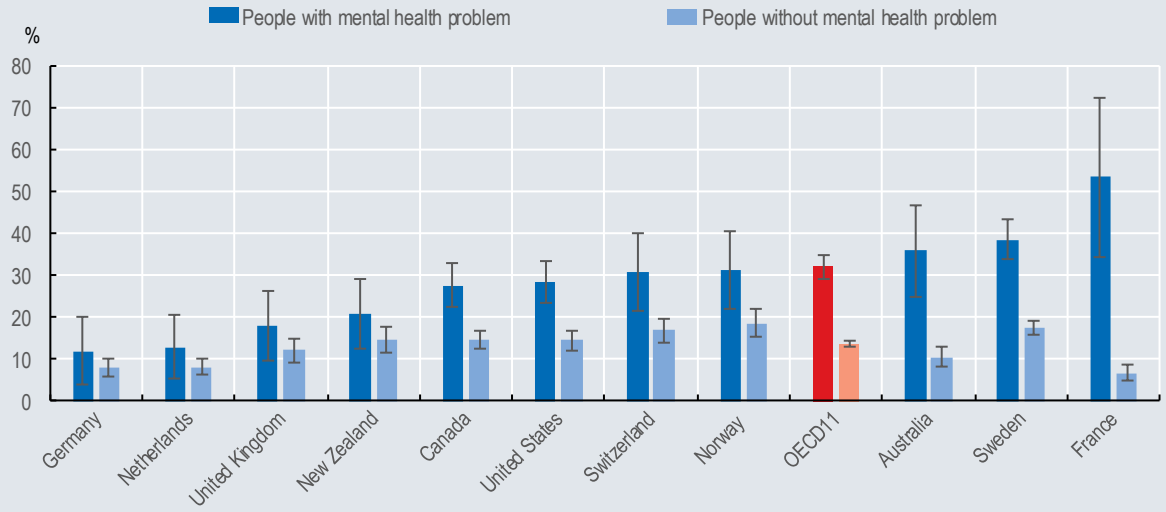
The Commonwealth Fund 2016 International Health Policy Survey of Adults (The Commonwealth Fund, 2016[45]) was conducted in 11 countries - Australia, Canada, France, Germany, the Netherlands, New Zealand, Norway, Sweden, Switzerland, the United

Kingdom, and the United States – with a total of 26 863 adults interviewed by phone about their experiences with their country's health care system, their health and well-being.

The survey included the question ***"Have you ever been told by a doctor that you have depression, anxiety or other mental health problems".*** Comparing responses across all the survey questions for respondents who answered 'yes' with those who responded 'no' to the mental health problem question can shed light on how people with a mental health problem in the participating countries navigate the health system.

Respondents who answered 'yes' reported similar experiences to the remaining respondents in some areas of care. In others, their reported care experience appears to be inferior. In several countries, for example, people with a mental health problem were significantly more likely to report have received conflicting information from different health care professionals (see Figure 3.10). The differences were most pronounced in Australia, Sweden and France.

Figure 3.10. People with a mental health problem are more likely to report conflicting information from health care professionals



Note: 'People with a mental health problem' are the respondents who answered "yes" to the question "thinking about the past 2 years, when receiving care for a medical problem, was there EVER a time when you received conflicting information from different doctors or health care professionals?" Data limitations. The number of respondents in the 11 countries ranged from 1 000 (Germany) to 7 124 (Sweden). The response rate ranged from 10.9% (Norway), 16.9% (Sweden) to 32.4% (Netherlands) to 46.9% (Switzerland). The sample sizes of respondents who answered 'yes' to the mental health question were therefore small in many cases. In addition, the mental health survey question does not permit distinguishing between individuals who were suffering from a mental health problem at the time of the survey, and those who had experienced mental ill-health in the past but have since recovered. Additionally cultural and linguistic differences in how the question was interpreted could influence responses. Results have not been risk-adjusted for co-morbidities and socio-economic status.
Source: OECD analysis based on Commonwealth Fund 2016 International Health Policy Survey (The Commonwealth Fund, 2016[45]).

3.1.15. Collaboration to enhance patient-reporting in mental health

Given the health and economic impact of mental ill-health, it is important to assess the quality and outcomes of care in this area. Existing outcome and process indicators – while very useful in some circumstances – do not provide the entire picture of quality and performance. This information gap impedes efforts to improve care, practice and policy.

However, patient-reporting in mental health is still at a relatively nascent stage. Data collection is patchy, and routine reporting and use of the information is far from the norm. As of 2018, only five of the twelve countries surveyed (Australia, Israel*, Netherlands, Sweden, United Kingdom – England) reported that PROMs and PREMs were collected regularly in the mental health setting. Only Australia, the Netherlands and England collected *and* routinely reported both. As such, a limited pool of national data exists that are not readily comparable at an international level.

This needs to change and the OECD has been working with patients, clinicians and policymakers and other experts from 13 countries to develop PREM and PROM data collection standards in mental health to enable international reporting, and foster the capacity to collect and use this important information in OECD countries.

* The statistical data for Israel are supplied by and under the responsibility of the relevant Israeli authorities. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law.

Chapter 4. Patient-reported indicators for the assessment of primary / ambulatory care for people with (multiple) chronic conditions

This chapter describes the key indicator domains to be used to measure outcomes of care and experiences with care of patients with (multiple) chronic conditions who are largely managed in primary care or other ambulatory care settings. For each indicator domain, a number of existing indicators are defined and several pros and cons of different instruments are discussed.

4.1. The value of the PaRIS survey to health systems and policy

The main policy objective of PaRIS is to contribute to developing more people-centred health systems by evaluating their relative performance in delivering people-centred care and in meeting the needs of patients, motivating necessary policy reforms and health care quality improvement initiatives. PaRIS will also encourage incorporating patient-reported outcomes and experiences as part of routine primary/ambulatory care, thus supporting improvements in people-centred care.

Within the PaRIS survey, a new set of internationally comparable measures will enable policy makers to monitor:

- Outcomes and experiences of patients with one or more chronic conditions nationwide;
- How outcomes and experiences develop over time;
- How country results compare to other countries; and
- How outcomes and experiences vary among specific groups of adults.

Furthermore, policy makers will be able to identify:

- Groups of adults that require specific attention; and
- Aspects of health care that require specific attention.

The PaRIS survey will help to answer the following key policy questions:

- What are the patient-reported outcomes of care for adults with chronic conditions and multi morbidity, such as overall health status, physical health status, mental health status and social health status?
- What are the experiences of care of adults with chronic conditions and multi morbidity such as experiences of care accessibility; quality of the communication between care providers; shared decision-making with care providers; care continuity and coordination; comprehensiveness of care; and safety and trust?
- How do outcomes and experiences vary for adults with chronic conditions by background characteristics such as age group, sex, education level, occupational status, household composition, health-risk behaviours, level of multi-morbidity, disease status and confidence in managing one's own care?
- How do key characteristics of ambulatory/primary care organisations relate to the outcomes and experiences of adults with chronic conditions?

The PaRIS survey data will provide for greater understanding of the quality of health care measured through the experiences and outcomes of patients at both national levels (within country) and at the international level (across countries). Further, it will help shed better light on the primary care sector. Surveys of health information infrastructure in OECD countries clearly describe the deficiency in national health data about the quality and performance of primary care. Fewer than half of OECD countries reported having national patient-level primary care data and only two regularly linked such data to outcomes to report on health care quality (OECD, 2015[20]).

The PaRIS survey will also offer other benefits to health systems. The outreach to and engagement with adults with chronic conditions, ambulatory/primary care service providers, policy makers and the public throughout the development and implementation of PaRIS will foster a broader dialogue about both the importance of people-centredness; and the feasibility and benefits of incorporating person-reported outcomes and experiences as part of routine primary/ambulatory care. This will further be explained in Chapter 3.

Policy goals that are outside of the scope of PaRIS are the use of benchmarks for public reporting of identifiable providers and the use of PaRIS results for performance-based reimbursement schemes and pay-for-performance arrangements. There are two reasons why this will not be possible. First, this would set an extremely high bar for the standardisation and validation of the data and methods that would be unfeasible in a pioneering project. Second, since results would have important consequences for providers, this would undermine their support of PaRIS and harm participation rates and data quality. This would most likely reduce the usefulness of the survey for strengthening patient and provider involvement and supporting health care improvement.

4.2. Survey indicators and analysis enabled by the PaRIS survey

Through indicators and in-depth analysis, the PaRIS survey will support understanding outcomes and experiences from the perspective of adults with chronic conditions and will point to potential areas where policy actions could improve outcomes and experiences.

The PaRIS survey data will be analysed to produce indicators of outcomes and experiences that are disaggregated by country and by key characteristics of patients, of ambulatory/primary care providers and of national health systems. This will be enabled by the structure of the PaRIS survey, which will consist of a questionnaire to be completed by surveyed adults with chronic conditions who are served by ambulatory/primary care service providers. The questionnaire measures outcomes and experiences of health care and provides background information about patient characteristics; and a supporting questionnaire directed to ambulatory/primary care service providers to provide some background information about structural characteristics that relate to adults' self-reported outcomes and experiences of care. Indicators will be featured in OECD publications, such as Health at a Glance, and in stand-alone publications of tables and figures comparing countries.

In addition, several related analyses will be possible: for example:

- analyses of variations in outcomes and experiences across and within countries;
- analysis relating PaRIS indicators to health system characteristics and other indicators collected by OECD;

- and analyses relating PaRIS indicators to information on policies to make health systems more people-centric as part of the Health Committee programme of work on people-centred health systems.

There would also be the possibility for countries to oversample to address their unique policy needs.

Box 4.1 provides an illustration of possible outputs they would be able to find from the PaRIS survey.

Box 4.1. Possible output of the PaRIS International Survey of Patients with Chronic Conditions

With the data collected in the PaRIS survey, many different questions can be answered and specific comparisons can be made. Examples of the information that the PaRIS survey would yield:

- Scores of self-rated health (scale 1 to 5) among diabetes patients with co-morbidities by country and / or individual background characteristics
- Scores on anxiety and symptoms of depression scale among people who were diagnosed with cancer in the past five years, by country
- % of patients with two or more chronic conditions that got a medication review in the previous year (review of all medication used) by country
- % of patients with two or more conditions that reported to have one provider who is coordinating all care, by country or region
- % of patients with dementia that attended an emergency department in the past six months, by country
- Confidence in managing one's own care among people with chronic conditions by country and age group
- Geographical differences within countries in the extent to which people with chronic conditions experience access problems.
- Scores on fatigue scale among patients with chronic heart conditions, variation within and between countries
- Trust in healthcare system among people with chronic conditions broken down by socio-economic status and country

4.3. Themes addressed by the PaRIS survey

The PaRIS survey will cover health outcomes and experiences of health care of adults with chronic conditions. This information will be both about the targets of the survey (the individual), and some contextual information would also be collected about his or her provider.

4.3.1. Adults with chronic conditions: key themes and indicators

Tables 4.1 and 4.2 present the main domains of indicators and other variables that will be covered in the PaRIS survey. The major content themes are self-reported socio-demographic characteristics, health behaviours and morbidity, health outcomes, confidence in managing one's own care, and health care experiences. Sub-domains of health care experiences include accessibility, communication, shared decision-making, continuity and coordination, comprehensiveness and safety and trust.

The choice of the domains, subdomains and indicators is based on both literature review and discussions with the international Taskforce. Previous version were reviewed by the member countries and adjusted based on their feedback. It will be necessary to test the length (completion time) of the survey to manage survey costs and minimise response burden.

Table 4.1. Domains and indicators for patient-reported outcomes and experiences for consideration in the PaRIS survey

THEME	INDICATOR DOMAIN	PATIENT-REPORTED INDICATORS
Patient-reported Health outcomes (PROMs)	**Overall health status**	**Health status summary score**
	Physical health status	**Physical health status summary score**
		Self-rated health status i.e. excellent, good, fair
		Pain intensity and interference with life activities
		Fatigue intensity and interference with life activities
		Physical function i.e. ability to walk/climb stairs
		Sleep quality rating and sleep disturbances
		Limitations in typical activities of daily life/disabilities
		Body functions
	Mental health status	**Mental health status summary score**
		Anxiety i.e. fearfulness and worries that overwhelm
		Symptoms of depression i.e. feelings of worthlessness,/hopelessness
	Social health status	**Social health status summary score**
		Ability to participate in social activities
		Social roles and responsibilities
		Social activities limitations i.e. inability to participate in activities
		Satisfaction with participation in social roles
		Satisfaction with participation in discretionary activities
Patient-reported experiences (PREMs)	**Overall experience**	**Experience summary score**
	Accessibility	Ease of appointment booking, waiting time,
	Communication	Time given for consultation, treated with respect, opportunity to ask questions
	Shared decision-making	Discussion of patient goals and priorities for their care, Patient involvement in developing the treatment plan
	Continuity and coordination	Frequency cared for by the same person, lab tests repeated because a carer hasn't access to test results, Pharmacist or PHC professional reviews all used medications with patient

Table 4.2. Patient level characteristics for consideration in the PaRIS survey

THEME	INDICATOR DOMAIN	PATIENT CHARACTERISTICS
Patient characteristics	**Patient engagement and activation**	**Patient Activation Summary Score**
		Believes active role is important
		Has confidence and knowledge
		Takes action
		Stays the course under stress
	Health literacy	Ability to understand written information pertaining to medical condition
		Confidence with Forms
		Required assistance when reading
	Socio-demographic	Age, sex, education level, occupational status, household composition, household size
	Health behaviors	Smoking status, BMI, physical activity, alcohol consumption
		Multi-morbidity level: Number of conditions
	Health conditions	Conditions:
		- Alzheimer's disease or other cause of dementia
		- Arthritis or ongoing problem with back or joints
		- Breathing condition such as asthma or COPD
		- Cancer (diagnosis or treatment in the last 5 years)
		- Diabetes type 1 or 2
		- Chronic kidney or liver disease
		- High blood pressure
		- Cardiovascular condition / heart condition
		- Depression
		-Other mental health condition such as bipolar disorder or schizophrenia
		- Neurological condition (such as epilepsy or migraine)
		- Another chronic condition
	Recent healthcare use	In the last 12 months, have you had any unexpected stays in hospital because of your condition (or conditions)?
		Have you visited an emergency/urgent care centre in the past 12 months? Was this related to one of your conditions?

4.3.2. Patient-reported Health Outcomes

PROMs for patients with chronic conditions form one of the core elements of the survey. PROMs are measures of health status or health-related quality of life, as reported by patients. They can broadly be categorised as either generic or condition-specific measures. Generic PROMs measure the wellbeing of all types of patients, regardless of their condition. Therefore, they can be applied to a wide range of patients and will be included in the PaRIS survey for patients with chronic conditions. Generic instruments seek to describe health and improvements in health in terms of the impact on health-related quality of life, broadly construed. The main domains covered by generic PROMs are general health, physical health function and symptoms, mental health function and symptoms, and social health (Tirli Bryan et al., 2014[46]). The PaRIS survey will include all these domains. These domains entail several subdomains, which will be part of PaRIS survey as well.

4.3.3. Generic Instruments to measure PROMs

There has been a proliferation of measures of quality of life and health status over the last two decades. In 2000, 1,275 separate measures existed, and the production of new measures was considerably growing (Garratt et al., 2002[47]). As such, there is a need for generic,

comparable measures. To inform the incorporation of generic instruments in the PaRIS survey, the OECD secretariat hosted a workshop in January 2017 with experts in the field to discuss the relevant instruments currently available. Amongst this variety of measures, previous studies have highlighted three instruments important for measuring PROMs: the SF-36®, PROMIS 10 and EQ-5D.

All three instruments provide a means of describing (generating a "profile" of) health. The SF-36® applies an algorithm to patients' responses to individual questionnaire items and scales to produce two summary scores: one for physical health and one for mental health. The EQ-5D uses a visual analogue scale to elicit from patients a single score for their overall health. The PROMIS 10 does not yield an overall score but gives physical health and mental health component scores.

Extensive research has been done on the validity, reliability, reproducibility, and utility of health status surveys when applied to general audiences and sub-groups based on age, sex, nationality, and disease entity. The EQ-5D and SF-36® have both been used in large surveys of the general public – population norms are available for both that may have relevance in "benchmarking" performance. Both generic measures – EQ-5D and SF-36® – have validated translations of their instruments available in a range of languages. The PROMIS instrument, being much newer, has a limited evidence base.

Table 4.3. Comparison of the three generic PROMs instruments

PROM instrument	Domains measured	No. of questions	Discrimination between domains and pt. types and populations	Validation	Translation
SF-36	9 domains: Physical functioning (PF) (10), Rolelimitation-Physical (RP) (4), Bodily Pain (BP) (2), General Health (GH) (5), Vitality (VT) (4), Social Functioning (SF) (2), Rolelimitation-Emotional (RE) (3), Mental Health (MH) (5); health transition (1)	36	Able to detect differences between groups defined by age, sex, socio-economic status, geographical region and clinical conditions	Validated across a range of conditions, settings and languages	170 languages
EQ-5D	4 domains: Anxiety/Depression (1), Mobility (1),Pain/Discomfort (1), Self-Care (1), UsualActivities (1)	5	underperforms in the Pain/Discomfort dimension	Validated in a diverse patient population in several countries, including patient groups with chronic conditions (cardiovascular disease, respiratory disease, depression, diabetes, liver disease, personality disorders, arthritis, stroke)	170 languages
PROMIS Global 10	5 domains: physical function, fatigue, pain, emotional distress, and social health	10	N/A	Newer than the previous two measures. Validated for some conditions and patient groups	8 languages

Source: Measuring Generic Patient Reported Outcome Measures for Health System Improvement, OECD, 2017. OECD Workshop on Generic PROMs.

The SF-36 is a widely used generic measure of health status. Thirty five of the 36 items are grouped into eight scales that address health constructs considered to be important to most health care situations: physical functioning, role limitations (physical problems), bodily pain, general health, vitality, social functioning, role limitations (emotional problems), and mental health. One item assesses perception of changes in health but is not used to compute scale scores. The SF-36v1® was developed in 1990 during the Medical Outcomes Study

(MOS) to measure generic health concepts relevant across age, disease, and treatment groups. It is available in 161 languages (Aaronson et al., 1992[48]).

The EQ-5D has two components: EQ-5D Visual Analogue Score asks patients how good or bad [their] health is today, on a scale of 0 (worst) to 100 (best). The second component, the EQ-5D index score, asks patients to indicate their current health status in dimensions of mobility, ability to undertake self-care, ability to undertake usual activities, pain/discomfort, and anxiety/depression. It was developed in 1990 and available in 170 languages.

The PROMIS Global Health questionnaire was developed by the National Institutes of Health in 2009. The questionnaire includes a 10-question survey that assesses generic health-related quality of life compared with population norms. PROMIS 10 gives a summary indicator of health status by assessing five domains: physical function, fatigue, pain, emotional distress, and social health. Nine of 10 questions are answered using 5-point Likert scales, and the 10th question is answered using a numerical rating scale.

4.3.4. Crosswalks between generic PROMs

A key issue related to the use of generic PROMs is comparability between different generic PROMs instruments. Currently, ggeneric health status is reported using multiple measures, versions, value sets. In order to establish internationally comparable data, common measures and metrics are needed. However, as the as the set of compared countries increases, choices for comparable PROMs diminish.

Various mapping algorithms (or crosswalks) from SF-12 and PROMIS-10 to the EQ-5D have been published in the literature and are currently available. However, these algorithms typically offer specific, delimited solutions. For example, they may only map to a particular version of EQ-5D (most map to the 3L measure), and while some map to health states, others map only to specific country valuations (e.g. UK only or US only value sets). Thus, the use of these crosswalks helps to establish a common comparison metric but offers a limited solution set for the types of international comparisons that can be made. Crosswalks also "compound" the error of score estimates, as they incorporate error from the original measure and the resultant measure. For these reasons, alignment on the use of one tool is recommended, with crosswalks to be used only when alignment cannot be achieved.

4.4. Overall health status

4.4.1. Self-rated health status

Self-rated health status (SRH) is a common measure in population health surveys. SRH is typically measured as a single-item, the most common wording of which is "In general, would you say your health is" with the response items "excellent," "very good," "good," "fair," or "poor" (Bombak, 2013[49]). The construct, and measure response options, can vary based on the country of analysis. The OECD collects data on this indicator, and to address differences in construct, groups self reported health into the following categories: 'good/very good' (or excellent) (all positive response categories), 'fair' (not good, not bad), 'bad/very bad' (all negative response categories). SRH has been reported to be significantly associated with various health related topics such as mortality, recovery after illness, and health service utilization (DeSalvo et al., 2006[50]). SRH may vary over time and can be contextual, as it reflects the personal evaluation of health. For example, people with poor mental health can distort perceptions of self-related health and vice versa. Self-rated health status has already been collected for international comparisons using general populations.

4.5. Physical health status

4.5.1. Pain intensity and interference with life activities

Pain is a common symptom associated with chronic illnesses such as arthritis, cancer, neurological conditions, among others. The biopsychosocial model of pain acknowledges that pain is the result of complex interaction among biological, psychological and social factors. This model also emphasizes the individual differences in pain perception deriving from differences in various cognitive, affective, behavioural and social factors (Covic et al., 2003[51]). A recent review of pain assessment in clinical care identified over 20 PROMs—both condition/procedure patient-centred specific and generic, were being used to assess patient levels of pain (Holmes et al., 2017[52]). Pain measurements can be visual scales, or single- or multi-item questionnaires. Single-item measures such as a Numeric Rating Scale (NRS) measure pain using a 0-10 pain scale, where 0 is 'no pain' and 10 is 'pain as bad as it can be'. Single item measures are easy for assessment because they assess key dimensions of pain, are quick to administer and easy to score. However, multi-item questionnaires such as the Brief Pain Inventory provide a comprehensive assessment of pain, by measuring not only pain intensity but also how it interferes with quality of life domains, addressing general activity, walking, mood, sleep, work, relations with other persons and enjoyment of life (Tan et al., 2004[53]). The EQ-5D includes a domain related to pain, asking respondents to rate themselves on the following scale: 1) I have no pain or discomfort, 2) I have moderate pain or discomfort, or 3) I have extreme pain or discomfort.

Research led by the private sector suggests that pain levels exist are high globally, with more than 60% of individuals experiencing bodily pain on a weekly basis in Australia, USA, Canada, Mexico, Spain, Russia, Philippines, Colombia, Portugal, and Romania (GSK, 2017[54]). However, pain levels appear to have high variance across countries. For example, a study of orthopaedic procedures across the US and 13 other countries found that US patients reported worse pain following procedures, which did not appear to be associated with surgery type or perioperative pain risk factors (Zaslansky, Meissner and Chapman, 2018[55]). Tools have also been validated for use in multiple countries. The American Pain Society Patient Outcome Questionnaire, for example, has been validated for use in Iceland, Australia, and Denmark (Zoëga, Ward and Gunnarsdottir, 2014[56]) (Botti et al., 2015[57]).

4.5.2. Fatigue

Fatigue is a common symptom associated with a wide range of chronic diseases. Measures of fatigue can be either generic or disease specific. The may be designed with the intention of designating the level of severity of fatigue or to classify individuals as either fatigued or non-fatigued. A study assessing the reliability and validity of PROMs related to fatigue found that almost 40 generic and disease-specific PROMs for assessing fatigue were currently in use (Nordin et al., 2016[58]). A related review did not identify any instrument that met all pre-established criteria for an ideal instrument—but noted that several instruments demonstrated good psychometric properties. In addition to psychometric properties, some tools demonstrated the ability to track changes over time—namely the Fatigue Severity Scale [FSS], the Brief Fatigue Inventory [BFI], the Fatigue Symptom Inventory [FSI], and the Multidimensional Assessment of Fatigue [MAF] (Whitehead, 2009[59]). Another tool is PROMIS Fatigue, a seven item scale, has been found to have good reliability and validity across diverse populations. Example items from the PROMS F-SF are: "How often did you feel tired," and "How often were you too tired to take a bath/shower". Response options are on a 5-point Likert scale, ranging from 1 = *never* to 5

= *always*. (Ameringer et al., 2016[60]) Research using the FSS has assessed levels of fatigue for patients with multiple sclerosis residing in over 50 countries or territories (Weiland et al., 2015[61]).

4.5.3. Sleep quantity rating and sleep disturbances

Research has demonstrated associations between sleep disturbances and obesity, diabetes, cardiovascular disease and mortality. People suffering from a chronic illness are at risk to develop sleeping disorders disturbances (Smagula et al., 2016[62]). Conversely, sleep disorders have shown to increase risk for cardiovascular and metabolic disorders (Grandner et al., 2012[63]). Tools to assess sleep quality have been used to identify sleep deprivation or disorders, as well as sleep quality and sleep disturbances. There are several existing tools to assess sleep quality, including the Insomnia Severity Index (7 items), the Epworth Sleepiness Scale (8 item questionnaire), and the Pittsburgh Sleep Quality Index (9 items) (Mastin, Bryson and Corwyn, 2006[64]). A wide variety of self-administered tools was identified in a recent systematic review of tools for sleep assessment (see Table 4.2) (Ibáñez, Silva and Cauli, 2018[65]). Current measures identify poor sleep quality survey respondents related various sleep related topics, for example if they have ever experienced difficulty falling asleep, difficulty maintaining sleep or consumption of sleeping pills. (Mesas et al., 2014[66]). Previous research studying sleep quality and HQoL in France, USA, and Japan has been conducted using the Insomnia Severity Index (Léger et al., 2012[67]). However, overall, there appears to be limited international research assessing sleep quality across a large number of countries.

Table 4.4. Self-administered questionnaires for the detection of sleep disorders

	Sleep questionnaire	Structure	Period
MSQ	Mini Sleep Questionnaire (Zoomer et al., 1985)	10 items (7 point scale)	Recently
PSQI	Pittsburgh Sleep Quality Index (Buysse et al., 1989)	9 items (4 point scale)	1 month
ESS	Epworth Sleepiness Scale (Johns, 1991)	8 items (4 point scale)	Recently
ISI	Insomnia Severity Index (Morin, 1993)	7 items (5 point scale)	Recently
SDQ	Sleep Disorders Questionnaire (Douglass et al., 1994)	175 items (5 point scale)	Recently
SACS	Sleep apnea clinical score (Flemons et al., 1994)	4 items (100 point scale)	Recently
FOSQ	Functional Outcomes of Sleep Questionnaire (Weaver et al., 1997)	30 items (4–5 point scale)	Recently
SAQLI	Calgary Sleep Apnea Quality of Life Index (Flemons & Reimer, 1998)	35 items (7 point scale)	1 month
OSQ	Oviedo Sleep Questionnaire (Bobes et al., 1998)	15 items (4–7 point scale)	1 month
BQ	Berlin Questionnaire (Netzer et al., 1999)	10 items (2–5 point scale)	Recently
ASQ	Athens Sleep Questionnaire (Soldatos, Dikeos & Paparrigopoulos, 2000)	8 items (4 point scale)	1 month
SEMSA	Self-efficacy in Sleep Apnea (Weaver et al., 2003)	26 items (4 point scale)	Recently/Future
SQ	STOP Questionnaire (Chung et al., 2008)	4 items (2 point scale)	Recently
SBQ	STOP-BANG Questionnaire (Pallesen et al., 2008)	8 items (2 point scale)	Recently
BIS	Bergen Insomnia Scale (Chasens, Ratcliffe & Weaver, 2009)	6 items (8 point scale)	1 month
FOSQ-10	Functional Outcomes of Sleep Questionnaire—10 (Takegami et al., 2009)	10 items (4 point scale)	Recently
SFV	Simple Four Variables (Chai-Coetzer et al., 2011)	4 items (2–6 point scale)	Recently
OSA50	Obesity, Snoring, Apneas, aged over 50 (Chai-Coetzer et al., 2011)	4 items (3–4 point scale)	Recently

Source: (Ibáñez, Silva and Cauli, 2018[65]).

4.5.4. Physical functioning – Body functions and limitations in daily life activities

Physical functioning refers to the capacity to perform activities, which require physical ability. It can be classified in both body functions and the basic activities of daily living (ADL), essential for self-care (Gross, Jones and Inouye, 2015[68]). Limitations in conducting the activities of daily life and altered body functionality are a paramount concern for individuals with chronic illness.

Functioning and impairment screens are methods to determine an individual's level of independence. For example, the Washington Group general disability measure was created with the aim of creating an internationally comparable measure of disability. The short version of the measure asks respondents about their ability seeing, hearing, walking, remembering, self-care, and communicating. For each of these questions, respondents can choose between the response categories of "No difficulty, some difficulty, a lot of difficulty, or unable to do it" (Palmer and Harley, 2012[69]). Similarly, the Activities of Daily Living (ADLs) functioning screen asks about the following activities: bathing, dressing, transferring, using the toilet, continence, and eating. Respondents can be asked to reply on a difficulty scale (e.g. no difficulty, some difficulty, a lot of difficulty, cannot do)

or on an assistance scale (Can perform without assistance, can perform with assistance, cannot perform (with assistance)). Previous studies compared physical functioning and ADL performance internationally. One of these studies resulted in the development of the International Physical Activity Questionnaire (IPAQ), which is developed in collaboration with 12 countries. It measures activity in daily life amongst four domains, transportation, at work, during household and gardening tasks and during leisure time (Hagströmer, Oja and Sjöström, n.d.[70]).

However, questionnaires can also include both items about limitations and difficulties performing ADL tasks and body functions to capture physical functioning. The PROMIS instrument incorporates 165 items in its Physical Functioning Domain, which measure the functioning of upper, lower and central extremities as well as activities of daily living. The EQ-5D includes questions about mobility, self-care and usual activities which are aligned with the construct of physical functioning (Van Reenen and Janssen, 2015[71]).

4.6. Mental health status

The two primary domains for PRO assessments of mental health are symptoms of anxiety and symptoms of depression.

Anxiety disorders are the most common type of mental health conditions and have a major impact on lives. Moreover, feelings of emotional distress are experienced privately and are therefore a good candidate when measuring PROMs. People suffering from anxiety are at increased risk of several health-related conditions, such as hypertension and coronary health disease (Roest et al., 2010[72]) Tools that assess anxiety are numerous. A systematic review identified more than 140 different scales used to measure anxiety, including generic tools and tools used to assess anxiety in response to particular situations in the medical field (Rose and Devine, 2014[73]). Anxiety is frequently measured by multi-item self-report questionnaires such as the Hamilton Anxiety Rating Scale and the Becks Anxiety Inventory, which consist out of 14 items and 21 items, respectively.

Depression is a both a symptom of and a chronic disease in itself. Moreover, depression has been associated with other chronic conditions, including heart disease and cancer (Caruso et al., 2017[74]). Research has found that clinical interviews can significantly under report rates of depression when compared to self-reported measures of depression (Lim et al., 2018[75]). There are several self-reported multi-item checklist to capture symptoms of depression. The Beck Depression Inventory is one of these and is a widely used self-report instrument for measuring symptoms of depression, and was translated in many languages (Dere et al., 2014[76]). A second example is the Centre for Epidemiologic Studies Depression Scale (CES-D 10), which is used in cross-national comparison studies (González et al., 2017[77]).

Prevalence estimates of depression symptoms have been found to vary across countries (Targum, Nakagawa and Sato, 2013[78]). This might reflect true differences in prevalence, but also measurement error because of methodological factors such as differences in item performance across settings. There has been several attempts to overcome cultural differences in measuring depression, by cross-validation of depression instruments (Wang et al., 2016[79]).

Symptoms of anxiety and depression are assessed independently but also conjointly. The Hospital Anxiety and Depression Scale is a 14-item self-report screening scale to assess possible presence of both anxiety and depressive states (SLOEKERS et al., 2002[80]). The EQ-5D also combines anxiety and depression into one item, asking respondents to rate

themselves on the following scale: 1) I am not anxious or depressed. 2) I am moderately anxious or depressed. 3) I am extremely anxious or depressed (Norman et al., 2009[81]). Similarly, SF-36 groups emotional problems in the following question: "During the past 4 weeks, have you had any of the following problems with your work or other regular daily activities as a result of any emotional problems (such as feeling depressed or anxious)" (Ware and Gandek, 1998[82]).

Anxiety and depression are also key domains in the PROMIS framework. Within the PROMIS framework, the concepts measured separately as two domains. The PROMIS tool for Anxiety contains 29 items, which measure self-reported fear, anxious misery, hyper arousal and somatic symptoms. The Depression Domain consist of 28 items, measuring self-reported negative mood, views of self, social cognition and decreased positive affect and engagement.

4.7. Social health status

Social health is a key component of overall well-being. Measures of social health have been tied with other PROs such as fatigue, pain, and other outcomes that can hinder an individuals ability to participate in their usual family and social roles. Measurements of social health have been developed for PROMIS—current tools include over 160 individual items related to social health (see Table 4.5). Items related to Social Health were developed with input from the PROMIS Social Health Workgroup, who dually developed a framework for conceptualizing social health. Item banks were created with the intent of collecting information on the social health status of both individuals with health conditions, as well as the experience of healthy people. (Hahn et al., 2010[83]). Common domains of social health status include measures to evaluate an individual's participation in social activities, social roles and responsibilities, satisfaction with participation in social roles, and satisfaction with participation in discretionary activities.

SF-36 contains one question related to social health status, asking, "During the past 4 weeks, how much of the time has your physical health or emotional problems interfered with your social activities (like visiting with friends, relatives, etc.)?" (Ware and Gandek, 1998[82]). Finally, EQ-5D does not include measures of social health. For some scales, items are constructed as written statements, using a 7-day reporting period. Ability items used in PROMIS use a 5-point frequency rating scale (Never, Rarely, Sometimes, Often and Always) and the Satisfaction items use a 5-point intensity rating scale (Not at all, A little bit, Somewhat, Quite a bit and Very much) (Hahn et al., 2010[83]).

Table 4.5. PROMIS Measures Related to Social Health

Adult Domains	Definition	Bank/ Scale/ Pool #items	Short Forms #items
Ability to Participate in Social Roles and Activities	Perceived ability to perform one's usual social roles and activities.	35	4, 6, 8
Companionship	Perceived availability of someone with whom to share enjoyable social activities such as visiting, talking, celebrations, etc.	6	4
Emotional Support	Perceived feelings of being cared for and valued as a person; having confidant relationships.	16	4, 6, 8
Informational Support	Perceived availability of helpful information or advice.	10	4, 6, 8
Instrumental Support	Perceived availability of assistance with material, cognitive or task performance.	11	4, 6, 8
Satisfaction with Participation in Discretionary Social Activities (v1.0)	Contentment with leisure interests and relationships with friends.	12	7
Satisfaction with Participation in Social Roles (v1.0)	Satisfaction with performing one's usual social roles and activities	14	4, 6, 7, 8
Satisfaction with Social Roles and Activities (v2.0)	Satisfaction with performing one's usual social roles and activities (e.g., "I am satisfied with my ability to participate in family activities").	44	4, 6, 8
Social Isolation	Perceptions of being avoided, excluded, detached, disconnected from, or unknown by, others.	14	4, 6, 8

Source: http://www.healthmeasures.net/explore-measurement-systems/promis/intro-to-promis/list-of-adult-measures

4.8. Patient-reported Experience Measures (PREMs)

Understanding a patient's experience when he or she receives health care is integral to improving people-centred care. There has been an increased recognition of the importance of the patient's perspective in providing quality health care. Capturing and reporting patient experience is an important part of the overall health system performance measurement efforts. Across countries, health service providers, administrators and policy-makers have indicated a desire and need for comparable patient-reported indicators to better understand and improve quality of health care service delivery and outcomes.

Over the last decade, increased attention has been placed on measuring and improving the health care experience of patients. The number of countries measuring patient experience has been increasing over time. Almost all OECD countries have at least one national survey measuring PREMs. In almost all of the OECD countries with national data collection, PREMs have been collected through national population-based surveys, and/or surveys covering patients who have had a recent experience either in an outpatient or inpatient care setting.

Table 4.6. List of OECD Patient Experience Indicators

Indicator Name
-Consultation skipped due to costs
-Medical tests, treatment or follow-up skipped due to costs
-Prescribed medicines skipped due to costs
-Waiting time of more than four weeks for getting an appointment with a specialist
-Patients reporting having spent enough time with any doctor during the consultation
-Patients reporting having spent enough time with their regular doctor during the consultation.
-Patients reporting having received easy-to-understand explanations by their regular doctor
-Patients reporting having had the opportunity to ask questions or raise concerns to any doctor
-Patients reporting having had the opportunity to ask questions or raise concerns to their regular doctor
-Patients reporting having been involved in decisions about care or treatment by any doctor

Source: https://stats.oecd.org/

The OECD has been measuring patient experience indicators related to access to health care, autonomy in care and treatment decisions and communication with physician during ambulatory care due to their relevance and importance across health systems (see Figure 4.6 for the list of indicators). These indicators were based on the work of the Commonwealth Fund and other national and international PREMs surveys. The set of indicators covers important aspects of people-centred care which are common across health systems such as patient participation/involvement, good relationship between the patient and health professionals including clear and open communication, and the appropriate context in which care is delivered including access to care (Kitson et al., 2013[84]). The domains of these OECD indicators, particularly communication and access to care, are considered to be important dimensions of health care quality (Mohammed et al., 2016[85]).

4.8.1. Accessibility

Patients with chronic conditions often require frequent and timely access to the health care system. There are several key domains related to accessibility as an indicator of quality, including appointment and provider accessibility, waiting time to get appointment or to reach the provider, as well as ease of making an appointment and communicating with the care provider (Mohammed et al., 2016[85]). Other access related topics may include domains related to payment and geographical accessibility. Current suggested indicators for accessibility include ease of appointment booking and waiting time.

Table 4.7. Possible Survey Questions for Accessibility

Accessibility Domain	Possible Survey Question
Ease of appointment booking	How easy or difficult is it to get medical care in the evenings, on weekends, or holidays without going to the hospital? (Commonwealth Fund, 2016[86])
Waiting time	Thinking about this last consultation, which of the following best describes the type of care you principally received? How quickly did you get an appointment to see this <healthcare provider>? Was the time you waited for the appointment a problem for you? (OECD, 2018) On the actual day of the consultation, how long did you wait (for example in the doctor's waiting room) before you were actually seen? Was the time you waited to be seen a problem for you? (OECD, 2018)

4.8.2. Communication

Clear and effective communication with care providers is essential for patients with chronic diseases, who often require consistent interaction with care providers and play a key role in their own health care. Moreover, good communication has been associated with improved health outcomes. Research studying patients with acute MI found that good communication was associated with better post-discharge health-related quality of life (correlation coefficient 0.33) (Hutchings, Varagunam and Black, 2013[87]). A review of 26 studies identifying attributes of high quality communication, found that dimensions could include the following: the provider's ability to listen, clearly explain conditions and treatments, display language concordance, and demonstrate good relationship/rapport with patients (Mohammed et al., 2016[85]). Current suggested indicators for communication that could be captured in PREMs include time given for consultation, treatment with respect, and the opportunity to ask questions.

Table 4.8. Possible Survey Questions for Communication

Communication Domain	Possible Survey Question
Time given for consultation	Did this <doctor/nurse/allied health professional> spend enough time with you? (OECD, 2018)
Treated with respect	During [the course of your care], how often did [your doctor/nurse/allied health professional] treat you with courtesy and respect? (Commonwealth Fund, 2016[86]) (adapted from source).
Opportunity to ask questions	Did this <doctor/nurse/allied health professional> give you an opportunity to ask questions or raise concerns about recommended treatment? (OECD, 2018)

4.8.3. Shared decision-making

Shared decision-making is an important mechanism for empowering patients to serve as active participants in their health care. PREMs in this area may capture whether or not a patient involvement in discussion of their care goals and priorities for their care, or their involvement in developing a care plan. A 2011 review of measures of shared decision making identified seven tools that solicit information on the shared decision making process from the patient perspective (Scholl et al., 2011[88]). Identified tools included items related to the patient's level of comfort in regard to the decision-making process and results, incorporation of patient preferences, and patients' self-confidence or belief in abilities for decision making, among other domains. Current suggested indicators of shared decision

making include discussion of patient goals and priorities for their care and patient involvement in developing the treatment plan.

Table 4.9. Possible Survey Questions for Shared Decision-making

Shared Decision-making Domain	Possible Survey Question
Discussion of patient goals and priorities for their care	During the past year, when you received care, has any health care professional you see for your [insert condition] discussed with you your main goals or priorities in caring for this condition (Commonwealth Fund, 2016[86]).
Patient involvement in developing the treatment plan	Did this <doctor/nurse/allied health professional> involve you as much as you wanted to be in decisions about your care and treatment? (OECD, 2018)

4.8.4. Continuity and care coordination

Similar to other PREM domains, continuity of care and care coordination are underpinning characteristics of high quality care for patients with chronic conditions. Research assessing the PROMs and PREMs for patients undergoing elective survey included an item on coordination of care as well as three items on if the patient received sufficient discharge information. The question, based on a question used in UK Inpatient Surveys run by the Picker Institute, states "Sometimes… a member of the staff will say one thing and anther will say something quite different. Did this happen to you? [Yes, often/Yes, sometimes/ No]" (Black, Varaganum and Hutchings, 2014[89]). This question could be adapted for use in primary care. Survey items used by the Commonwealth Fund include the following question related to care coordination from the perspective of primary care, "How often does your regular doctor or someone in your doctor's practice help coordinate or arrange the care you receive from other doctors and places?" (Commonwealth Fund, 2016[86]). Similarly, the Patient Assessment of Chronic Illness Care (PACIC) measure assessing patient perceptions of chronic illness care includes a sub-scale on follow-up/care coordination (Glasgow et al., 2005[90]). Suggested domains for patients with chronic conditions include items related to the frequency by which the patient is cared for by the same person, lab tests repeated because a carer has no access to test results, and that a pharmacist or PHC professional reviews all used medications with patient.

Table 4.10. Possible Survey Questions for Continuity and Care Coordination

Continuity and Care Coordination Domain	Possible Question
Frequency cared for by the same person	Is there one doctor you usually go to for your medical care? (Commonwealth Fund, 2016[86]) Not counting any time you may have been hospitalized, how many different doctors have you seen in the past 12 months? (Commonwealth Fund, 2016[86])
Lab tests repeated because a carer hasn't access to test results	Now thinking about the past 2 years, when receiving care for a medical problem, was there EVER a time when doctors ordered a medical test that you felt was unnecessary because the test had already been done? (Commonwealth Fund, 2016[86])
Pharmacist or PHC professional reviews all used medications with patient	In the past 12 months, has a doctor or pharmacist reviewed with you all the medications you take? (Commonwealth Fund, 2016[86])

4.9. Health literacy

Health literacy is defined as an individual's knowledge, motivation and skills to access, understand, evaluate and apply health information. Health literacy is both an input and an outcome from the interaction between patients and health systems. It depends on the patient's skills and ability to understand and apply knowledge about health (Moreira, 2018[91]). However, professionals may use medical jargon, drug instructions are not always clear, and health information in clinical settings can be complex and challenging to navigate. Research has shown that better health literacy is related to better health care outcomes.

Health literacy measures may include multiple components such as health knowledge, oral understanding and navigation skills. Existing tools to examine health literacy appear to be limited. The recently developed Health Literacy Questionnaire (HLQ) assesses multiple domains of health literacy. The tool has nine scales and a total of 44 items in total. Items are scored from 1-4 in some of the scales (Strongly Disagree, Disagree, Agree, Strongly Agree), and from 1-5 in others (Cannot Do, Very Difficult, Quite Difficult, Easy, Very Easy) (Hawkins et al., 2017[92]) Other existing tools to assess health literacy include the National Assessment of Adult Literacy (NAAL), the Functional Health Literacy in Adults (TOFHLA) and the Rapid Estimate of Adult Literacy in Medicine (REALM).

HL has been measured in 18 OECD countries using general skills surveys (e.g. PIAAC) or HL-specific surveys (e.g. European Health Literacy survey). The 2012 European Health Literacy Survey (EU-HLS) allowed, for the first time, the international comparison of HL levels. Respondents answered 47 questions on *accessing, understanding, evaluating and applying* health information according to three health domains: health care, disease prevention and health promotion. Respondents answered questions following a scale that ranged between very difficult to very easy. Answers were aggregated into four levels of HL: *insufficient, problematic, sufficient and excellent. Insufficient* and *problematic* levels are considered *limited* HL. Since 2012, the EU-HLS has become a benchmark for HL measurement across various OECD countries.

Although the EU-HILS has been a ground-breaking tool, there is an ongoing debate on what aspects to include when measuring Health literacy. Several available tools contain much more items than could be included in the PaRIS questionnaire, so a pragmatic and short selection should be applied. A possible candidate is a short three-item questionnaire to assess functional health literacy on the personal level; the SBSQ (Chew, Bradley and Boyko, 2004[93]) (Chew et al., 2008[94]). The SBSQ consists of three items: 'How confident

are you filling out forms by yourself?' (confidence with forms), 'How often do you have someone (like a family member, friend, hospital/clinic worker or caregiver) help you read hospital materials?' (assistance required when reading) and 'How often do you have problems learning about your medical condition because of difficulty understanding written information?' (understanding written information). Responses were scored on a 5-point Likert scale ranging from 0 (always/not at all confident) to 4 (never/extremely confident).

4.10. Patient Engagement and Activation

Tools that capture domains related to patient engagement and activation can give policy makers insight on patient's abilities and perceived roles in managing their own health—an important domain for patients with chronic conditions. Common domains of patient-reported confidence include aspects of confidence with forms, ability to self-manage care, ability to manage stress, and ability to understand and learn about relevant health conditions. Measures of patient activation have been used encourage self-management of health conditions by ensuring that care provided is appropriately matched to a patient's skills, motivation and confidence, while also considering the individual's needs and capabilities (Hibbard and Gilburt, 2014[95]). Furthermore, such indicators have been found to be associated with better health outcomes in primary care (Mattingly and Nong, 2019[96]). Additional research suggests that patient activation may be a more accurate predictor of health outcomes socio-demographic factors such as ethnicity and age (Hibbard and Gilburt, 2014[95]). As in other domains, there is a rich existing environment of measures on the topic. A 2015 systematic review identified 19 different tools to assess patient empowerment, Six generic measures, and 13 for patients with specific conditions including diabetes, cancer, and mental illnesses (Barr et al., 2015[97]). Each of these tools were assessed to demonstrate their inclusion of four key domains: patient states experiences and capacities, patient actions and behaviours, and patient self-determination within the healthcare relationship, and the development of patient skills.

Table 4.11. Existing Measures of Patient Engagement and their Respective Domains

Measure (target population)	Domain 1 Patient states, experiences and capacities	Domain 2 Patient actions and behaviours	Domain 3 Patient self-determination within the healthcare relationship	Domain 4 Developing patient skills
*Patient Empowerment Scale # 1 [13] (All patients, generic)	+	-	-	-
Kim Alliance Scale [46,47] (All patients, generic)	+	+	+	-
Treatment Related-Empowerment Scale (TES) [45] (All patients, generic)	+	-	+	-
Health Education Impact Questionnaire [44] (All patients, generic)	+	+	-	+
*Scale developed by Bann et al. (2010) [49] (All patients, generic)	+	+	+	+
Health Care Empowerment Inventory (HCEI) [52] (All patients, generic)	+	-	+	-
Empowerment Scales (Original) [28–31] (Mental health patients)	+	+	-	-
Empowerment Scales (Version 2) [36] (Mental health patients)	+	+	-	-
Health Promotion Intervention Questionnaire [32,33] (Mental health patients)	-	-	+	-
*Empowerment Questionnaire for Inpatients (EQuIP) [37] (Mental health patients)	+	-	+	-
Consumer Evaluation Of Mental Health Services (CEO-MHS) Scale [34] (Mental health patients)	+	-	-	-
Inpatient Consumer Survey (Mental health patients)	-	-	+	-
Diabetes Empowerment Scale (Original) [16,41,42] (Diabetes patients)	+	+	-	-
Diabetes Empowerment Scale (Version 2) [39,40] (Diabetes patients)	+	+	-	-
Diabetes Empowerment Scale (Version 3) [38] (Diabetes patients)	+	+	-	-
Chinese Diabetes Empowerment Process Scale (C DEPS) [43] (Diabetes patients)	+	-	+	-
*Patient Empowerment Scale #2 [17] (Cancer patients)	+	-	-	-
*Cyber Info-Decisional Empowerment Scale (CIDES) [53] (Cancer patients)	+	-	+	-
*Genetic Counseling Outcome Scale [18] (Clinical Genetics patients)	+	-	-	-
Parents' Postnatal Sense of Security (PPSS) [48] (Postnatal patients)	+	-	+	-
Psoriasis Empowerment Enquiry in the Routine practice questionnaire (PEER) [50] (Psoriasis patients)	+	-	-	+
The Swedish Rheumatic Disease Empowerment Scale (SWE-RES-23) [51] (Rheumatology patients)	+	-	-	-

Source: (Barr et al., 2015[97])

A commonly used tool is the Patient Activation Measure (PAM), which has been found be associated patient outcomes for patients with diabetes (Remmers et al., 2009[98]) care. The PAM tool has been used in various countries, including the US, the UK and the Netherlands (Rademakers et al., 2012[99]). Items and response categories included in the PAM can be found in Figure 4.5.

Table 4.12. Items Included in the Patient Activation Measure

1.	When all is said and done, I am the person who is responsible for taking care of my health	Disagree Strongly	Disagree	Agree	Agree Strongly	N/A
2.	Taking an active role in my own health care is the most important thing that affects my health	Disagree Strongly	Disagree	Agree	Agree Strongly	N/A
3.	I know what each of my prescribed medications do	Disagree Strongly	Disagree	Agree	Agree Strongly	N/A
4.	I am confident that I can tell whether I need to go to the doctor or whether I can take care of a health problem myself.	Disagree Strongly	Disagree	Agree	Agree Strongly	N/A
5.	I am confident that I can tell a doctor concerns I have even when he or she does not ask.	Disagree Strongly	Disagree	Agree	Agree Strongly	N/A
6.	I am confident that I can follow through on medical treatments I may need to do at home	Disagree Strongly	Disagree	Agree	Agree Strongly	N/A
7.	I have been able to maintain (keep up with) lifestyle changes, like eating right or exercising	Disagree Strongly	Disagree	Agree	Agree Strongly	N/A
8.	I know how to prevent problems with my health	Disagree Strongly	Disagree	Agree	Agree Strongly	N/A
9.	I am confident I can figure out solutions when new problems arise with my health.	Disagree Strongly	Disagree	Agree	Agree Strongly	N/A
10.	I am confident that I can maintain lifestyle changes, like eating right and exercising, even during times of stress.	Disagree Strongly	Disagree	Agree	Agree Strongly	N/A

Source: (Hibbard, 2008[100])

Understanding the vital role of patients in providing input on this area of measurement, in particular, PaRIS initiative will work with a patient advisory group to receive input on this domain, including the domain name, sub-domains, potential indicators and items.

4.11. Primary/ambulatory care settings: Key themes and indicators

Collecting information on structural characteristics of primary care providers can help with understanding cross-country differences in the way care for adults with chronic conditions is organised and possible consequences for outcomes and experiences. Table 4.2 illustrates the structural characteristics that could be collected from the primary/ambulatory care providers who are caring for the adults with chronic conditions who will be participating in PaRIS.

Table 4.13. Potential primary/ambulatory care settings characteristics

THEME	SUB-THEME	CHARACTERISTICS
PHC Provider-reported characteristics	Clinic Characteristics	Rural/urban clinic location Postal Code (for geolocation) Structure of the care setting (e.g. solo practitioners, care team, integrated care provider, HMO) Staff composition of the provider (physicians, nurses, nutritionists etc.) Patient load
	Main PHC Professional's Characteristics	Age of patient's main PHC professional Professional designation of patient's main PHC professional (GP, Geriatrician, Specialist, nurse, etc.) Remuneration method of patient's main PHC professional Accreditation/certification status of the patient's main PHC professional

There is an enormous variation in primary/ambulatory care settings both across and within countries. In most cases, patients receive care from several individual care providers. These providers may have different configurations, for instance they can work together in small independent partnerships or in bigger organisations such as HMO's. Services can be contracted out by the main provider; they can cooperate in networks or be completely independent from each other.

Despite the variety of organisational arrangements, a common set of key characteristics, such as staff composition and patient load can be collected from all countries. Furthermore, countries interested in specific settings could add organisation-related variables to the short list of variables collected in all countries.

Chapter 5. Sampling strategy for the international survey of patients with chronic conditions

This chapter describes the sampling methodology for the survey of patients with chronic conditions. More specifically, it will discuss the definitions of the population of eligible patients and providers, how a multi-stages sampling will be implemented and how providers and patients will be recruited in such a way that satisfactory response rates will be realised.

5.1. A multi-staged survey of community dwelling adults with chronic conditions

The PaRIS survey focuses on people with one or more chronic conditions who live in private homes and whose conditions are being managed in primary care or other ambulatory care settings. A multilevel sampling design will be used. This will pose additional challenges to sampling and analyses of data compared to population based surveys. The advantage of this multilevel approach is that it will enable exploration in levels of variation across providers, health regions or other meaningful units and, in the longer term, to identify factors on different levels that correlate with better outcomes. The eligible population, sampling techniques, and strategies to recruit providers will be further clarified in this section.

5.2. The population of interest

As was noted previously, the PaRIS survey focuses on adults with one or more chronic conditions who are receiving health care from primary/ambulatory care service providers. Such adults are managing their conditions at home and may experience difficulties with care coordination and communication across multiple health care service providers.

These adults are, therefore, dwelling in private homes (community dwelling) and are not residing within institutions, such as nursing homes, homes for disabled persons, schools, prisons, hospitals, or any other institutionalised residential facility. In future phases, opportunities to expand the focus to institutions may be considered. However, in this stage including institutionalised people will significantly complicate the survey, as there is significant variation in the design and use of special care arrangements within and between countries.

Indicators will be reported for adults with chronic conditions overall and for key sub-groups of adults, such as:

- adults with multiple conditions;
- adults with specific conditions of interest, such as diabetes or heart disease;
- adults within certain age ranges, such as those who are aged 65 and older;
- adults with certain health behaviours, such as smoking status and BMI; and
- adults with certain socio-economic characteristics, such as education level and employment status.

5.2.1. Criteria for the definition of eligible population

The PaRIS Taskforce discussed the pros and cons of potential definitions of eligible survey participants. Three overarching recommendations emerged.

The first recommendation was to define the eligible population in a manner that enables primary/ambulatory care service providers to participate in the survey *regardless of the sophistication of their medical record keeping*. That is, providers with paper-based records (or electronic medical records that are not coded) should be able to participate. In so doing, all countries can participate in the survey and a bias due to a method that favours providers with electronic medical records or specific coding practices will be avoided.

The second was to *maximise the ease of participation in the survey* for both providers and adults with chronic conditions. Therefore, from the provider perspective, the method of selecting eligible adult participants should not be time consuming or difficult.

The third was to *ensure that the sample of participating adults is inclusive* of adults who do not visit their provider often while also minimising recall error from visits too long in the past.

Based on discussions with experts, it was concluded that selecting patients who visited the physician within three months of the survey would be too short a time period (too few patients who see their provider infrequently).Selecting patients who visited within the previous 12 months would be too long a time period (too likely to select patients whose last visit was too long ago to remember it clearly). Therefore 6 months seems to be an optimum.

5.2.2. Definition of the eligible population

Following the recommendations above, the patient population eligible for participation in the survey was defined as follows:

- Community-dwelling adults who are aged 45 and older AND
- Who visited their primary/ambulatory care provider at least once within the six months preceding the survey.

The above definition of the eligible population is inclusive of adults who do not have any chronic conditions. However, from national studies in many OECD countries it is known that the prevalence of chronic conditions increases with age and that by age 65, two-thirds or more of the population has at least one chronic condition From previous studies in Canada, Netherlands and Scotland, 40-50% of all people (including children) have at least one chronic condition. We can be confident that at ages 45 and older, the proportion of the total population with at least one chronic condition would lie between 40% and 66% (Ward and Black, 2016[101]; Barnett et al., 2012[102]; Koné Pefoyo et al., 2015[103]). Further, within the population aged 45 and older who also sought care from an ambulatory/primary care provider, the proportion with at least one chronic condition would, reasonably, be even higher.

In the patient questionnaire, patients will be asked to indicate whether he or she has one or more chronic conditions out of a list, and will be asked to write-in any other chronic conditions.

Serious conditions emerge at mid-life more frequently among people with less education and lower incomes. Selecting the population aged 45 and older will enable measurement

of important differences in outcomes and experiences by education level and other social or economic factors that will help with understanding how to improve people-centred care for vulnerable sub-groups.

5.2.3. Health care providers included in the study

Across the OECD, there has been a clear trend in recent decades to shift chronic care management and routine care toward generalists working in ambulatory/primary care settings. This trend accelerated in most countries in the past decade (Thomson et al., 2014[104]; Watkins, 2014[105]). Even in countries without a strong primary care tradition (e.g. Germany, Belgium), most of the population visits an ambulatory/primary care provider regularly. For example, a German study found that among patients who had a myocardial infarction, almost 86% visited a GP at least once (5.8 times on average) in the 12 months after the myocardial infarction. Of the remaining part, most (83%) visited an internist (Pohl et al., 2017).

The eligibility criteria for the ambulatory/primary care settings to be included in the survey were discussed extensively with experts and were refined based on these discussions. These eligibility criteria have already been established for international primary care accreditation, such as the *Joint Commission International Accreditation Eligibility Criteria, Primary Care*. The OECD will work together with participating countries to identify the ambulatory/primary care providers that provide services to community-dwelling patients that are similar to those described in the eligibility criteria, such as providing care to a defined community and providing services that address chronic disease management.

The exact providers eligible for participation in the study will be defined together with experts within each participating country. Inevitably, there will be differences across countries, particularly between countries with a strong primary care model and those where ambulatory/primary care is more fragmented or where the care arrangements may differ depending on patient factors, such as insurance status.

Primary/ambulatory care settings include single-handed practitioners and practitioners working in an ambulatory/primary care clinic with a team of health care professionals. These practitioners may be generalist doctors, such as general practitioners and family physicians, and in some countries they may also be nurses, internists, geriatricians and other health professionals who deliver routine care to patients with chronic conditions and coordinate their care with providers outside their clinic. Appropriate definition and inclusion criteria for ambulatory/primary care settings will be developed in each participating country to allow for the selection of comparable patient samples.

5.3. A multi-staged sample

A multi-stage sampling approach allows sampling adults from within ambulatory/primary care settings. Alignment with national directions and initiatives is a key principle of PaRIS. Therefore, each country's sampling plan may be adapted somewhat to national or local characteristics and information requirements. Some countries already have national surveys in place that enable reaching a representative sample of patients with chronic conditions and linking them to their ambulatory/primary care provider. In some cases, creating synergy with such existing surveys may be more efficient than launching a new survey, as long as the data collected are comparable.

5.3.1. Stage 1: Selection of areas/communities

In the first stage, a representative sample of areas within participating countries will be identified. The selection of areas/regions for participation in the study facilitates the engagement and participation of primary/ambulatory care providers and the adults whom they serve, which will help achieve necessary response rates.

This sampling method has been used in previous surveys, such as the US National Health and Nutrition Examination Survey (NHANES) and the Canadian Health Measures Survey (CHMS); and has enabled local-area focussed media, stakeholder and participant engagement and communications strategies to yield good response rates.

Advantages for PaRIS in selecting areas in the first stage of the sample design include the following:

- All countries can participate, even if there are no national registries of ambulatory/primary care providers and the ambulatory/primary care provider landscape is fragmented across multiple private and public care systems. There may be registries of providers at the small area/community level or there may be simple and feasible ways to work with the community to enumerate all of the ambulatory/primary care providers that are serving that community.
- Areas can be identified many months before the PaRIS survey will be implemented, enabling local area engagement and communications activities to be initiated well in advance of data collection.
- Medical and other community leaders can be engaged to support the promotion of PaRIS in their community.
- Community events to introduce PaRIS to health care providers and to the public in the local areas can be hosted, such as meetings with ambulatory/primary care practitioners.
- A local coordinator for PaRIS can be recruited who has a trusted reputation and a good network in the area and can engage with local area ambulatory/primary care providers to introduce the survey and to encourage participation.

Some countries may be able to achieve the same level of community engagement while sampling from the whole population (rather than from selected areas). This may be the case in countries with highly centralised systems, with existing representative data collection networks and/or countries that are relatively small. An appropriate sampling strategy for each participating country will be adapted by the international survey team in conjunction with the NPM.

5.3.2. Selection of primary/ambulatory care providers within areas

In the second stage, eligible primary/ambulatory care providers will be sampled from within selected small areas. Often, ambulatory/primary care providers have a high workload and are overloaded with requests for participation in research. Participation in the PaRIS survey asks time and energy and may somewhat interfere with daily routines. Therefore, it can be expected that incentives will be required to convince providers to participate and, therefore, to secure an acceptable response rate. Low response rates are unacceptable because they can lead to an insufficient patient sample size and create a bias towards providers who are more open to research and/or the objectives of PaRIS.

Previous experiences within the Quality and Costs in Primary Care study (QUALICOPC) that had a comparable sampling design and which was carried out in 35 countries have yielded valuable insights that can be used in the PaRIS project. Groenewegen et al. (Groenewegen, Greß and Schäfer, 2016[106]) analysed the variation in response rates across countries and explained this by the different types of success factors and incentives.

This study showed that:

- There is little evidence for the effectiveness of financial incentives directed to providers;
- Participation rates were particularly high in small countries such as Iceland and Malta;
- Sharing data with providers, for analysis or benchmarking seems to be an effective incentive;
- The engagement of influential, trusted professionals increases participation; and
- The effectiveness of specific incentives differs from country to country.

The finding that financial incentives seem not to be very effective to motivate providers to participate in research is in line with other studies that were mainly focussed on the recruitment of doctors for clinical trials, although proper reimbursement for time spent appeared important (Draper et al., 2009[107]). Non-financial motivators that have been reported to be effective are professional obligation, possible medical care benefits, professional acknowledgement for participation (such as accreditation points) and personal acquaintance with the researcher (Draper et al., 2009[107]; Foy et al., 2003[108]; De Wit et al., 2001[109]).

The study by Groenewegen et al. (Groenewegen, Greß and Schäfer, 2016[106]) stressed the important role of local coordinators (in this proposal we use the term National Project Managers). These are typically people or institutions who are well-known and trusted within the professional community or who represent a professional organisation with authority in the community. The advantage of such local or national coordinators was also confirmed in previous OECD studies such as PIAAC and PISA. An explanation for the high participation rates found in small countries may be that personal acquaintance plays a role; with providers approached for a study by someone they know and trust, rather than by an organisation that is far away from their everyday practice.

Based on these experiences, the following strategies will be followed to stimulate participation among primary/ambulatory care settings:

- National and, where necessary, local Project Managers will be identified in all countries.
- A tailor-made recruitment strategy will be developed together with the national and local project managers, taking into account the national and local circumstances.
- Professional organisations, payers and other stakeholders, as appropriate, will be engaged in the survey and in the broader mission to make systems more people centred and to promote the use of patient-reported indicators.

- The data collection strategy will be focussed on minimising extra work and disruption of practice routines.
- Following an opt-in procedure and within the limits of privacy regulation, primary/ambulatory care settings participants will be offered access to aggregated data in which they can compare outcomes with peers.
- Data will be collected using a user-friendly interface.
- The value added of the use of patient-reported indicators both from a policy, as well as a quality improvement perspective, will be used as an argument to encourage participation, rather than financial incentives.

Please note that in some countries, national databases of primary/ambulatory care patients exist. In these countries it may be unnecessary to sample patients via providers and both patients and their providers can be approached directly.

5.3.3. Stage 3: Selection of eligible adults receiving ambulatory/primary care

A sample of eligible adults (see section 5.2.2) will be drawn from the population cared for by the primary/ambulatory care providers selected in stage 2. This method has two main advantages:

- First primary/ambulatory care providers are involved in the study and, therefore, they can use the study results for quality improvement. From the patients' perspective, this implies that patients not only provide survey data, but will also be contributing to their own care process by sharing valuable information with their health care practitioner. Results that are shared with providers will be aggregated (anonymously).
- Second, response rates among adults are likely to be higher when approached via their primary/ambulatory care provider with whom they have trusting relationship, than if approached directly by an institute with whom they are unfamiliar.

An engagement and communications strategy to promote patient participation in the survey will be developed with input from a patient advisory panel. The communications strategy may include various outreach materials, such as invitation letters and brochures, a website with information about the survey, and community information events, among other methods.

5.4. Survey questionnaires

Two questionnaires are planned for the PaRIS survey. The first questionnaire will be completed by participating adults with chronic conditions. All of the information sought regarding these adults' characteristics, health care experiences and health care outcomes will be self-reported by them via a patient questionnaire. No patient clinical data from patient records will be requested of or provided by clinicians for this survey. The questionnaire should take between twenty and thirty minutes on average for patients to complete.

The second questionnaire will be directed to ambulatory/primary care providers. Providers will complete a questionnaire seeking information about the structural characteristics of their practice/clinic.

Phase 1 work will, as appropriate, draw from existing instruments and be coordinated with national or international partners who have extensive experience measuring patient-reported outcomes and experiences. This will ensure that expertise in countries is leveraged, and duplications of tasks and costs are avoided. Examples of instruments that could contribute to PaRIS include PROMIS from Northwestern University in the United States (created to assess symptoms and functions in any condition), EQ 5D, SF36;[1] and the newly launched effort at ICHOM to develop a standard set to assess adults' overall health. There are also efforts by the Commonwealth Fund and the Picker Institute Europe to measure patient's health care experiences.

Information about chronic conditions and outcomes and experiences of patients will be self-reported by participating adults. The questionnaire will include a screening question with a list of the most prevalent chronic conditions in order to help participants to identify conditions. Participants will also be invited to write-in any other chronic condition they may have that is not on the list and to provide an indication of when their chronic conditions were first diagnosed.

The PaRIS survey will not assess outcomes and experiences that are specific to a single condition (for example, cancer patients' experience of chemotherapy). Rather, the PaRIS survey will ask people about aspects of care that are common to most patients with chronic conditions, such as the extent to which care is being coordinated, to what extent they are able to participate in aspects of life that are important to them, their experience of pain, quality of sleep, social functioning, etc.

5.5. Limitations and mitigations of the survey design

The design of the survey focuses on understanding the experiences and outcomes of health care of adults with chronic conditions. The survey focuses on people who received care at least once in the six months preceding the survey. Although experiences with accessibility will be measured, the survey is not optimised to understand the circumstances and health of adults who, for reasons of cost, distance, provider shortages or other reasons, do not access primary/ambulatory care providers and/or who have undiagnosed chronic conditions. By sampling patients who did receive care, people who are facing access problems will by definition be underrepresented. Population-based survey are a better option to understand the magnitude of access problems on system level. Information on access problems and unmet needs is available from national general population surveys and international population surveys, such as EU-SILC in Europe and this information is routinely reported within the OECD health indicators.

Information about access barriers and unmet needs from other sources should be used to place the PaRIS survey findings into an appropriate context. Further, the PaRIS survey must include variables necessary for appropriate risk/case-mix adjustment.

[1] EQ-5D: 5 dimensional instrument for generic health from EuroQual group; SF-36: the Short Form instrument to measure patient health consisting of 36 items;

In the survey design and development of instruments a balance is required between 'broad', an approach that is relevant for a big share of the population and for large patient groups in all countries, and 'deep', covering specific and more detailed types of outcome measures that are relevant for people with specific conditions. The PaRIS survey is 'broad' rather than 'deep'. PaRIS is innovative because it will focus on needs that are relevant for people with multiple conditions (such as integration and coordination of care and people-centredness of services). It will ask about aspects that are relevant for a variety of patient groups within our populations, such as socio-economic groups. Results will enable policy makers and other stakeholders to gain a broader national and international picture of outcomes and experiences; but the broader focus will likely make the PaRIS survey instruments less useful for enhancing clinical management.

To obtain information about the conditions that people have, the survey will rely mainly on self-reporting. In addition to consulting users of the health system if the system is serving their needs, there are numerous other important reasons to do consult patients directly, such as improved privacy and data protection, lower requirements for data systems and coding, the avoidance of a diagnosis registration bias and the reduction of burden and level of complexity for healthcare providers.

A disadvantage of this approach may be that this information could be more subjective and less specific than information coming from clinical registries. To minimise this bias, the PaRIS survey will build on previously successful survey methods of this kind. The list of chronic health conditions will be formulated in a way that would be understandable for most lay-people. It can, however, not be ruled out that conditions may be somewhat over- or underreported.

Chapter 6. Conclusions and next steps

6.1. PaRIS contributes to the transition towards more people-centred and knowledge based health systems

While the concept of health-related quality of life (QoL) has existed for almost three decades, it is not measured or reported systematically. Performance metrics in health tend to focus principally on inputs and outputs. Outcomes such as life expectancy are important, but they are silent on a range of other patient preferences, including those related to pain, function, QoL, and the experience of care itself. This means that the picture of healthcare and health system performance is missing an essential component.

This is why during the ministerial meeting in January 2017, Ministers of health asked the OECD to develop new ways to measure health system performance ─ the Patient-Reported Indicators Surveys (PaRIS). This effort is working to provide cross-country comparisons of patients' own experience of medical care and health care outcomes, so that policymakers, providers, and patients will better understand how health systems make a difference to people's lives. This is part of the next generation of health reforms; a transition towards more people-centred and knowledge based health systems.

6.2. PaRIS addresses both care for specific conditions as well as primary care for people with (multiple) chronic conditions

With the support of the European Commission, the OECD has commenced activities in the framework of PaRIS in 2017. The PaRIS survey is building international capacity to measure and compare care outcomes and experiences as reported by patients, using indicators that enable comparisons across countries. PaRIS is also helping to steer patient-reported indicators to evolve in a common direction internationally, enabling shared learning, development and research. Since the current state of the art in the areas of patient reported measures varies highly across sectors, PaRIS follows two streams of work in parallel:

- In areas where patient-reported indicators such as PROMs and PREMs already exist, the first work stream supports countries to accelerate the adoption and reporting of validated, standardised, internationally comparable patient-reported indicators.
- To address the need to understand the outcomes and experiences of people with one or more chronic conditions, the second work stream is to develop a new international survey. This focuses on adults with one or more chronic conditions who are receiving primary/ambulatory care services.

Since the use of patient reported measures for a number of conditions is already developed in several countries across the OECD, this work follows an approach that coordinates and steers existing efforts on the collection of patient-reported indicators. For patients with chronic conditions who live in the community and who are mainly served in primary care settings, there is hardly any existing systematic data collection. Here, the PaRIS work is developing an international survey from the ground up.

This report describes early results from the work that has been done within the framework of PaRIS in the first two years. Since PaRIS is a comprehensive, large-scale international initiative, with much innovative and pioneering work, it will take several years and the results described in this report should be seen as preliminary. The results as reported are based on in-depth research and many intensive interactions with a range of world-leading experts. They provide a solid foundation to build further on this important initiative.

More specifically, this report describes the following main elements of PaRIS:

- patient-reported indicators that PaRIS recommends to be adopted for specific conditions, in particular breast cancer, hip or knee surgery and mental health conditions;
- early international patient-reported data on these conditions;
- a framework with domains for health system performance indicators reported by patients with chronic conditions;
- a sampling methodology to survey these patients and their health care providers.

6.3. Patient-reported indicators to be adopted for specific conditions

6.3.1. Hip- and knee replacements:

Each year, over 2.2 million people undergo an elective hip or knee replacement in OECD countries. Knee replacement rates have doubled since the year 2000, while hip replacements have increased by 30%. Inter- and intra-country variation in rates can be as high as 5-fold (OECD, 2014[3]).

For patients undergoing hip or knee replacements, two types of instruments are being used: instruments to measure generic domains and condition-specific instruments. For the measurement of generic domains, the EQ-5D-3L index was chosen as the common metric because (a) the majority of countries use this instrument; (b) algorithms exist to convert – or map – scores from other generic instruments to the EQ-5D-3L. Score conversions were conducted using patient-level data. Since it has been proved that crosswalks can be made across different instruments, registries using other instruments, such as the SF-12 could also participate in the international data collection. EQ-5D is broadly used for the calculation of health utility scores (QALYs) and is available in most languages. However, a number new generic instruments are emerging that can also generate QALYs and that enable newer functionalities such as computer-adaptive testing (CAT).

The most common condition-specific instruments currently used are the ‘HOOS’ (and its variants) and the Oxford Hip Score for patients with hip replacements and ‘KOOS’ (and its variants) and the Oxford Knee Score for people with knee replacements. No algorithms to convert scores are available at present.

It should be noted that a new generation instruments are emerging, such as the PROMIS suite, which enable modular application to specific characteristics as well as CAT functionalities. They combine the sensitivity and precision of condition-specific instruments with the generalisability of the generic ones because results are expressed on a common scale/index. Using these new instruments is being considered by a growing number of healthcare systems. However, a key challenge is to convince the clinical

community and registry custodians of the value of changing from the more established instruments.

Preliminary data indicate that:

- In each country both procedures improved the pain, function and health-related QoL as reported by patients.
- A hip replacement at age 65 generated, on average, an additional 4.3 QALYs compared to an average of 3.3 QALYs for a knee replacement.
- Inter-country variation was observed but not pronounced, suggesting that the data collection and analysis were sound.

Breast Cancer:

Scales used in this area and used in OECD's first international data collection are the postoperative breast satisfaction scale of the breast conserving therapy (lumpectomy) and breast reconstruction modules of the Breast Q tool, an internationally validated instrument used to measure breast surgery outcomes reported by patients.

The breast satisfaction scales measure body image in terms of a woman's satisfaction with her breasts, about breast appearance (e.g., size, symmetry, softness), clothing issues (e.g., how bras fit; being able to wear fitted clothes) and location and appearance of scars. There are separate modules for lumpectomies, mastectomies and reconstructions, with each module consisting of multiple separate scales covering such issues as psychosocial wellbeing, sexual wellbeing, physical wellbeing, satisfaction with breasts and satisfaction with care. There are also implant-specific items, including the amount of rippling that can be seen or felt.

The scores from each scale of the BCT and Reconstruction scales, along with the other Breast Q scales can be transformed to an Equivalent Rasch Transformed Score of 1-100 to allow direct comparison between scales.

Ten sites spanning seven countries participated in a pilot collection of patient-reported outcomes data for women undergoing breast cancer treatment. The results suggest that:

- Comparative international analysis is possible but further work to refine indicators and conduct risk-adjustment is needed.
- Some interventions result in better patient-reported outcomes compared to others, but variation in care treatment modalities is observed.
- Information derived from patient-reported data can be very useful for women when deciding on the optimal treatment for their individual needs and preferences, and for payers and policymakers about the comparative cost-effectiveness and cost-utility of various treatments.

6.3.2. Mental health conditions

The structural measurement of patient-reported indicators is in its infancy. Compared to the other two types of conditions reported above, very little data is available and only a handful of countries have longer-term initiatives in this area. The CMWF health policy survey has yielded some insides in mental health care users' experiences with care. Despite several limitations, new analyses conducted on data from the 2016 Commonwealth Fund survey of 11 countries suggest that people with a mental health problem:

- report a worse care experience than those without mental health problems in some domains of health care;
- are more likely to report receiving conflicting information from health professionals suggesting that these patients are more likely to encounter fragmented care delivery.

For policy makers, these results suggest that countries may be falling short of providing people-centered care for patients who are managing a mental health problem. Given existing evidence and knowledge in this area, improving the care experience for people with mental health issues can contribute towards better clinical and social outcomes for these patients.

Overall, the results demonstrate that presenting valid and comparable results from patient-reported indicators at international level is eminently possible. However, capacity within and among countries must be increased to collect and report these data in a consistent and harmonised way. International data collection guidelines will be required. These will need to be developed in partnership with national and international stakeholders, including patients and health care professionals.

6.4. A framework for the measurement of patient-reported indicators among patients with chronic conditions

The PaRIS survey's patient-reported indicators are the core of the initiative. However, a number of other variables on the level of patients, health care settings and system level will be incorporated. These variables will enable us to 1) to put variation in these outcomes into context 2) to observe differences across groups of patients, care settings and geographical regions and 3) to make case-mix corrections.

PROMs used in the survey will focus on physical, mental and social functioning and overall health. Examples of more specific indicator domains are pain, fatigue, mobility, and social roles and responsibilities. The PROMs in the survey are often closely related, which could result in ambiguity between PROM definitions and which psychometric instruments best capture the underlying constructs. For example, the measurement of physical functioning is often measured as one umbrella concept. However, it might also be divided in subcategories such as body functions and limitations in daily activities. Another domain refers to mental health. The survey will focus on both anxiety and symptoms of depression, which can be measured separately or conjointly. During future work a decision will be made how the PROMs will be measured and which psychometric instruments will be used, while taking into account the information received from the experts in the field.

Examples of PREMs domains that will be used are communication, continuity and coordination and shared decision-making.

Patient level characteristics that will be captured include demographic characteristics (such as sex and age), epidemiological characteristics (such as the health conditions and the number of health conditions), health care utilization, and items related to the patient's ability to manage their own health. The latter contains two types of variables: measures about knowledge, motivation and skills to access, understand, evaluate and apply health information (health literacy) and measures of patient engagement and activation.

On the provider level, examples of suggested variables are demographic characteristics, structure characteristics of the care setting (solo practitioners, health centres, HMO, etc.), and information on staff composition. Eventually individual characteristic of the main

provider may be incorporated into the survey. Since care models highly vary across countries and even within countries, countries will have the opportunity to add provider-level variables tailored to the national or regional context. This way, the PaRIS survey could evaluate the relative performance of various care models.

6.5. Sampling methodology for the survey of patients with chronic conditions

The sampling design of the PaRIS survey is innovative and will generate different information than would be possible with a population-based survey. A multi-stage sampling approach allows sampling adults from within ambulatory/primary care settings. This complex approach has been informed by an international expert taskforce who has assisted the project in identifying solutions for a number of challenges. Modification for national or local characteristics may be required.

Stage 1: Selection of areas/communities

In the first stage, a representative sample of areas within participating countries will be identified. The selection of areas/regions for participation in the study facilitates the engagement and participation of primary/ambulatory care providers and the adults whom they serve, which will help achieve necessary response rates.

The selection of areas will include the following efforts and considerations to ensure the most representative sample:

- Work with community leaders to enumerate all of the ambulatory/primary care providers that are serving that community, in cases where there are no useable registries or lists of ambulatory/primary care providers
- Engagement with health care providers in the PaRIS-survey through networking, community events, recruitment of a local-area coordinator to represent the survey

Stage 2: Selection of primary/ambulatory care providers within areas

In the second stage, eligible primary/ambulatory care providers will be sampled from within selected small areas. The recruitment of primary/ambulatory care providers will require a tailor-made strategy developed in consultation with the NPM, with input from providers/ambulatory advisors.

The selection of primary/ambulatory care providers within areas will include the following efforts and considerations to ensure the most representative sample:

- Recruitment of local-area coordinators who are known and respected to assist with the recruitment of sampled providers
- Engagement with professional organisations, payers and other stakeholders, as appropriate to promote the survey and the broader mission to make systems more people centred and to promote the use of patient-reported measures.
- A data collection strategy that minimises extra work and disruption of practice routines.
- A user-friendly interface for data collection and information sharing with providers.

- Following an opt-in procedure and within the limits of privacy regulation, offering providers access to aggregated data in which they can compare their outcomes with area, national and international benchmarks.

Stage 3: Selection of eligible adults receiving ambulatory/primary care

A sample of eligible adults will be drawn from the population cared for by the primary/ambulatory care providers selected in stage 2. The recruitment of selected adults will require a tailor-made strategy developed in consultation with the National Project Manager and with input from an international patient-advisory panel and may include:

- Invitation letters and brochures
- A website with information about the survey
- Community information/events

An example of these challenges is the variation across countries in how providers and patients are registered. Not all countries have national registrations from which samples can be drawn. Similarly, only some countries have list systems, which means that drawing samples from patients who have visited a provider is not always straightforward. Another challenge relates to the recruitment of providers.

In order to tackle these challenges, national and local expertise and networks are crucial. Participating countries will appoint National Project Managers and will engage with care providers in an early stage. The Secretariat, together with an external contractor and National Project Managers, will develop tailor-made implementation plans on a country by country basis. Previous stratified studies in primary care settings, such as the English GP-Patient survey and the Quality and Costs in Primary Care (QUALICOPC) have yielded valuable and practical insights that are being used to inform PaRIS' work.

6.6. Next steps

The work described in this report has been carried out in the first two and a half years after the PaRIS initiative was launched. The two work streams of PaRIS are different in terms of scope as well as working process. The three working groups are currently in different stages of the process and will continue their attempts to create internationally comparable datasets and to provide directions in the measurement of PROMs and PREMs internationally.

In the coming years, the secretariat intends to expand the number of working groups, as well as the number of participating countries per working groups.

The survey of patients with chronic conditions follows three phases: a development phase, a field trial and the implementation of the main survey. The first cycle is expected to end in 2023, when the OECD will publish the first batch of international data.

Indicators and methodologies used in both work streams will be further refined and developed with the help of technical experts, country officials, networks of patients and other relevant stakeholders.

References

(n.a.) (n.d.), *PROMIS ADULT PROFILE INSTRUMENTS A brief guide to the PROMIS® Profile instruments for adult respondents*. [120]

(n.a.) (n.d.), *PROMIS User Manual (Version 1.1)*, http://www.nihpromis.org. [119]

Aaronson, N. et al. (1992), *International quality of life assessment (IQOLA) project The international Quality of Life Assesment (IQOLA)*, https://link.springer.com/content/pdf/10.1007/BF00434949.pdf (accessed on 18 July 2019). [48]

Abahussin, A. et al. (2019), "PROMs for Pain in Adult Cancer Patients: A Systematic Review of Measurement Properties", *Pain Practice*, Vol. 19/1, pp. 93-117, http://dx.doi.org/10.1111/papr.12711. [4]

Ameringer, S. et al. (2016), "Psychometric Evaluation of the Patient-Reported Outcomes Measurement Information System Fatigue-Short Form Across Diverse Populations.", *Nursing research*, Vol. 65/4, pp. 279-89, http://dx.doi.org/10.1097/NNR.0000000000000162. [60]

Australian Government - National Mental Health Commission (2016), *Media Release: The Impact of Poor Mental Health: An Economic Issue,*, http://www.mentalhealthcommission.gov.au/media/181670/Economics%20of%20Mental%20Health%20in%20Australia%20MR%20(3).pdf. [41]

Barnett, K. et al. (2012), "Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study", *The Lancet*, Vol. 380/9836, pp. 37-43. [102]

Barr, P. et al. (2015), "Assessment of Patient Empowerment - A Systematic Review of Measures", *PLOS ONE*, Vol. 10/5, p. e0126553, http://dx.doi.org/10.1371/journal.pone.0126553. [97]

Basch, E. (2017), "Patient-Reported Outcomes — Harnessing Patients' Voices to Improve Clinical Care", *New England Journal of Medicine*, Vol. 376/2, pp. 105-108, http://dx.doi.org/10.1056/NEJMp1611252. [13]

Basch, E. et al. (2017), "Overall Survival Results of a Trial Assessing Patient-Reported Outcomes for Symptom Monitoring During Routine Cancer Treatment", *JAMA*, Vol. 318/2, p. 197, http://dx.doi.org/10.1001/jama.2017.7156. [14]

Black, N. (2013), "Patient reported outcome measures could help transform healthcare.", *BMJ (Clinical research ed.)*, Vol. 346, p. f167, http://dx.doi.org/10.1136/bmj.f167. [17]

Black, N., M. Varaganum and A. Hutchings (2014), "Relationship between patient reported experience (PREMs) and patient reported outcomes (PROMs) in elective surgery", *BMJ Quality & Safety*, Vol. 23/7, pp. 534-542, http://dx.doi.org/10.1136/BMJQS-2013-002707. [89]

Bombak, A. (2013), "Self-Rated Health and Public Health: A Critical Perspective", *Frontiers in Public Health*, Vol. 1, p. 15, http://dx.doi.org/10.3389/fpubh.2013.00015. [49]

Botti, M. et al. (2015), "Cross-Cultural Examination of the Structure of the Revised American Pain Society Patient Outcome Questionnaire (APS-POQ-R)", *The Journal of Pain*, Vol. 16/8, pp. 727-740, http://dx.doi.org/10.1016/j.jpain.2015.03.016. [57]

Bray, F. et al. (2018), "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries", *CA: A Cancer Journal for Clinicians*, Vol. 68/6, pp. 394-424, http://dx.doi.org/10.3322/caac.21492. [28]

Bryan, S. et al. (2013), *What are the most effective ways to measure patient health outcomes of primary health care integration through PROM (Patient Reported Outcome Measurement) instruments?*. [112]

Calvert, M., D. O'Connor and E. Basch (2019), "Harnessing the patient voice in real-world evidence: the essential role of patient-reported outcomes", *Nature Reviews Drug Discovery*, http://dx.doi.org/10.1038/d41573-019-00088-7. [16]

Cameron, D., J. Ubels and F. Norström (2018), "On what basis are medical cost-effectiveness thresholds set? Clashing opinions and an absence of data: a systematic review", *Global Health Action*, Vol. 11/1, p. 1447828, http://dx.doi.org/10.1080/16549716.2018.1447828. [37]

Cardoso, F. et al. (2019), "Early breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up", *Annals of Oncology*, http://dx.doi.org/10.1093/annonc/mdz173. [29]

Caruso, R. et al. (2017), *Depressive spectrum disorders in cancer: prevalence, risk factors and screening for depression: a critical review*, http://dx.doi.org/10.1080/0284186X.2016.1266090. [74]

Chew, L., K. Bradley and E. Boyko (2004), "Brief questions to identify patients with inadequate health literacy.", *Family medicine*, Vol. 36/8, pp. 588-94, http://www.ncbi.nlm.nih.gov/pubmed/15343421 (accessed on 30 July 2019). [93]

Chew, L. et al. (2008), "Validation of Screening Questions for Limited Health Literacy in a Large VA Outpatient Population", *Journal of General Internal Medicine*, Vol. 23/5, pp. 561-566, http://dx.doi.org/10.1007/s11606-008-0520-5. [94]

Commonwealth Fund (2016), *2016 Commonwealth Fund International Health Policy Survey*, https://www.commonwealthfund.org/sites/default/files/documents/___media_files_publications_in_the_literature_2016_nov_1915_osborn_2016_intl_survey_ha_11_16_2016_questionnaire.pdf (accessed on 24 July 2019). [86]

Covic, T. et al. (2003), "A biopsychosocial model of pain and depression in rheumatoid arthritis: A 12-month longitudinal study", *Rheumatology*, http://dx.doi.org/10.1093/rheumatology/keg369. [51]

De Wit, N. et al. (2001), "Participation and successful patient recruitment in primary care", *Journal of Family Practice*, Vol. 50, pp. 976-981. [109]

Dere, J. et al. (2014), "Psychological Assessment Cross-Cultural Examination of Measurement Invariance of the Beck Depression Inventory-II", http://dx.doi.org/10.1037/pas0000026. [76]

DeSalvo, K. et al. (2006), *Mortality prediction with a single general self-rated health question: A meta-analysis*, http://dx.doi.org/10.1111/j.1525-1497.2005.00291.x. [50]

Desomer, A. et al. (2018), *Use of patient-reported outcome and experience measures in patient care and policy – Short report*, http://www.kce.fgov.be (accessed on 28 June 2019). [110]

Devlin, N. (2019), *Can We Really Compare and Aggregate PRO Data Between People and Settings? Implications for Multi-Country Clinical Trials and HTA*. [23]

Donovan, J. et al. (2016), "Patient-Reported Outcomes after Monitoring, Surgery, or Radiotherapy for Prostate Cancer", *New England Journal of Medicine*, Vol. 375/15, pp. 1425-1437, http://dx.doi.org/10.1056/NEJMoa1606221. [6]

Doty, M. et al. (2017), *Mirror, Mirror 2017: International Comparison Reflects Flaws and Opportunities for Better U.S. Health Care*, http://dx.doi.org/10.15868/socialsector.27698. [18]

Draper, H. et al. (2009), "Offering payments, reimbursement and incentives to patients and family doctors to encourage participation in research", *Family Practice*, Vol. 26/3, pp. 231-238. [107]

Foy, R. et al. (2003), "How evidence based are recruitment strategies to randomised controlled trials in primary care? Experience from seven studies.", *Family Practice*, Vol. 20, pp. 83-92. [108]

Fujisawa, R. and N. Klazinga (2018), *Health Working Papers OECD Health Working Papers No. 102 MEASURING PATIENT EXPERIENCES (PREMS): PROGRESS MADE BY THE OECD AND ITS MEMBER COUNTRIES BETWEEN 2006 AND 2016*, http://www.oecd.org/els/health-systems/health-working-papers.htm (accessed on 10 July 2019). [113]

Garratt, A. et al. (2002), "Quality of life measurement: bibliographic study of patient assessed health outcome measures", *BMJ*, Vol. 324/7351, p. 1417, http://dx.doi.org/10.1136/BMJ.324.7351.1417. [47]

Garratt, A. et al. (n.d.), *Papers Quality of life measurement: bibliographic study of patient assessed health outcome measures*. [111]

Glasgow, R. et al. (2005), "Development and validation of the Patient Assessment of Chronic Illness Care (PACIC).", *Medical care*, Vol. 43/5, pp. 436-44, http://www.ncbi.nlm.nih.gov/pubmed/15838407 (accessed on 11 July 2019). [90]

González, P. et al. (2017), "Measurement Properties of the Center for Epidemiologic Studies Depression Scale (CES-D 10): Findings from HCHS/SOL HHS Public Access", *Psychol Assess*, Vol. 29/4, pp. 372-381, http://dx.doi.org/10.1037/pas0000330. [77]

Grandner, M. et al. (2012), "Sleep disturbance is associated with cardiovascular and metabolic disorders", *Journal of Sleep Research*, http://dx.doi.org/10.1111/j.1365-2869.2011.00990.x. [63]

Greenhalgh, J. et al. (2018), "How do aggregated patient-reported outcome measures data stimulate health care improvement? A realist synthesis", *Journal of Health Services Research & Policy*, Vol. 23/1, pp. 57-65, http://dx.doi.org/10.1177/1355819617740925. [15]

Groenewegen, P., S. Greß and W. Schäfer (2016), "General Practitioners' Participation in a Large, Multicountry Combined General Practitioner-Patient Survey: Recruitment Procedures and Participation Rate", *International Journal of Family Medicine* 4929432. [106]

Groenewegen, P. et al. (2005), "What is important in evaluating health care quality? An international comparison of user views", *BMC Health Services Research*, Vol. 5/1, p. 16, http://dx.doi.org/10.1186/1472-6963-5-16. [125]

Gross, A., R. Jones and S. Inouye (2015), "Development of an Expanded Measure of Physical Functioning for Older Persons in Epidemiologic Research", *Research on Aging*, Vol. 37/7, pp. 671-694, http://dx.doi.org/10.1177/0164027514550834. [68]

GSK (2017), *GSK Global Pain Index 2017 Global Research Report GPI 2017 Global Report Findings 2 GPI 2017 Global Report Findings*, https://www.gsk.com/media/3814/global-pain-index-2017-report.pdf (accessed on 18 July 2019). [54]

Hagströmer, M., P. Oja and M. Sjöström (n.d.), "The International Physical Activity Questionnaire (IPAQ): a study of concurrent and construct validity", http://dx.doi.org/10.1079/PHN2005898. [70]

Hahn, E. et al. (2010), "Measuring social health in the patient-reported outcomes measurement information system (PROMIS): item bank development and testing", *Quality of Life Research*, Vol. 19/7, pp. 1035-1044, http://dx.doi.org/10.1007/s11136-010-9654-0. [83]

Hamdy, F. et al. (2016), "10-Year Outcomes after Monitoring, Surgery, or Radiotherapy for Localized Prostate Cancer", *New England Journal of Medicine*, Vol. 375/15, pp. 1415-1424, http://dx.doi.org/10.1056/NEJMoa1606220. [5]

Haroz, E. et al. (2016), "Depression symptoms across cultures: an IRT analysis of standard depression symptoms using data from eight countries", *Social Psychiatry and Psychiatric Epidemiology*, http://dx.doi.org/10.1007/s00127-016-1218-3. [121]

Hawkins, M. et al. (2017), "The Health Literacy Questionnaire (HLQ) at the patient-clinician interface: a qualitative study of what patients and clinicians mean by their HLQ scores", *BMC Health Services Research*, Vol. 17/1, p. 309, http://dx.doi.org/10.1186/s12913-017-2254-8. [92]

Hibbard, J. (2008), *Can Increasing Patient Activation Improve the Outcomes of Care?*, https://www.pcpcc.org/sites/default/files/media/Hibbard%20-%20Feb%2028_0_0_0.pdf (accessed on 24 July 2019). [100]

Hibbard, J. and H. Gilburt (2014), *Supporting people to manage their health: An introduction to patient activation*, https://www.kingsfund.org.uk/sites/default/files/field/field_publication_file/supporting-people-manage-health-patient-activation-may14.pdf (accessed on 24 July 2019). [95]

Holmes, M. et al. (2017), "The impact of patient-reported outcome measures in clinical practice for pain: a systematic review.", *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation*, Vol. 26/2, pp. 245-257, http://dx.doi.org/10.1007/s11136-016-1449-5. [52]

Hutchings, A., M. Varagunam and N. Black (2013), *Relationship between patient reported experience (PREMs) and patient reported outcomes (PROMs) in elective surgery*, http://www.lshtm.ac.uk (accessed on 11 July 2019). [87]

Ibáñez, V., J. Silva and O. Cauli (2018), "A survey on sleep assessment methods.", *PeerJ*, Vol. 6, p. e4849, http://dx.doi.org/10.7717/peerj.4849. [65]

ICHCAHPS (n.d.), *In-Center Hemodialysis CAHPS Survey: Survey and Protocols*, https://ichcahps.org/SurveyandProtocols.aspx (accessed on 10 July 2019). [114]

Institute for Health Metrics and Evaluation (2019), , http://www.healthdata.org/. [38]

Kern, P. et al. (2015), "Impact of Age, Obesity and Smoking on Patient Satisfaction with Breast Implant Surgery – A Unicentric Analysis of 318 Implant Reconstructions after Mastectomy", *Geburtshilfe und Frauenheilkunde*, Vol. 75/06, pp. 597-604, http://dx.doi.org/10.1055/s-0035-1546171. [32]

Kitson, A. et al. (2013), "What are the core elements of patient-centred care? A narrative review and synthesis of the literature from health policy, medicine and nursing", *Journal of Advanced Nursing*, Vol. 69/1, pp. 4-15, http://dx.doi.org/10.1111/j.1365-2648.2012.06064.x. [84]

Koné Pefoyo, A. et al. (2015), "The increasing burden and complexity of multimorbidity", *BMC Public Health*, Vol. 15/415. [103]

Konopka, J. et al. (2018), "Quality-Adjusted Life Years After Hip and Knee Arthroplasty", *JBJS Open Access*, Vol. 3/3, p. e0007, http://dx.doi.org/10.2106/jbjs.oa.18.00007. [27]

Léger, D. et al. (2012), "Chronic insomnia, quality-of-life, and utility scores: Comparison with good sleepers in a cross-sectional international survey", *Sleep Medicine*, Vol. 13/1, pp. 43-51, http://dx.doi.org/10.1016/J.SLEEP.2011.03.020. [67]

Le, Q. (2013), "Probabilistic mapping of the health status measure SF-12 onto the health utility measure EQ-5D using the US-population-based scoring models", *Quality of Life Research*, Vol. 23/2, pp. 459-466, http://dx.doi.org/10.1007/s11136-013-0517-3. [26]

Lim, G. et al. (2018), "Prevalence of Depression in the Community from 30 Countries between 1994 and 2014.", *Scientific reports*, Vol. 8/1, p. 2861, http://dx.doi.org/10.1038/s41598-018-21243-x. [75]

Luxford, K., D. Safran and T. Delbanco (2011), "Promoting patient-centered care: a qualitative study of facilitators and barriers in healthcare organizations with a reputation for improving the patient experience", *International Journal for Quality in Health Care*, Vol. 23/5, pp. 510-515, http://dx.doi.org/10.1093/intqhc/mzr024. [10]

Maas, C. and J. Hox (2005), "Sufficient Sample Sizes for Multilevel Modeling", *Methodology*, Vol. 1/3, pp. 85-91, http://dx.doi.org/10.1027/1614-2241.1.3.85. [123]

Mastin, D., J. Bryson and R. Corwyn (2006), "Assessment of Sleep Hygiene Using the Sleep Hygiene Index", *Journal of Behavioral Medicine*, Vol. 29/3, pp. 223-227, http://dx.doi.org/10.1007/s10865-006-9047-6. [64]

Matros, E. et al. (2015), "Cost-Effectiveness Analysis of Implants versus Autologous Perforator Flaps Using the BREAST-Q", *Plastic and Reconstructive Surgery*, Vol. 135/4, pp. 937-946, http://dx.doi.org/10.1097/prs.0000000000001134. [33]

Mattingly, T. and K. Nong (2019), "Implementing a Health Confidence Tool at Time of Discharge", *The Patient - Patient-Centered Outcomes Research*, Vol. 12/1, pp. 163-165, http://dx.doi.org/10.1007/s40271-018-0343-3. [96]

Medic, G., M. Wille and M. Hemels (2017), "Nature and Science of Sleep Short-and long-term health consequences of sleep disruption", http://dx.doi.org/10.2147/NSS.S134864. [118]

Mental Health Commission of Canada (2012), *Making the Case for Investing in Mental Health in Canada*, https://www.mentalhealthcommission.ca/sites/default/files/2016-06/Investing_in_Mental_Health_FINAL_Version_ENG.pdf. [44]

Mesas, A. et al. (2014), "Sleep quality and the metabolic syndrome: the role of sleep duration and lifestyle", *Diabetes/Metabolism Research and Reviews*, Vol. 30/3, pp. 222-231, http://dx.doi.org/10.1002/dmrr.2480. [66]

Mohammed, K. et al. (2016), "Creating a Patient-Centered Health Care Delivery System", *American Journal of Medical Quality*, Vol. 31/1, pp. 12-21, http://dx.doi.org/10.1177/1062860614545124. [85]

Moreira, L. (2018), *HEALTH LITERACY FOR PEOPLE-CENTRED CARE Where do OECD countries stand?*, http://www.oecd.org/els/health-systems/health-working-papers.htm (accessed on 25 July 2019). [91]

Nag, N. et al. (2018), "Development of Indicators to Assess Quality of Care for Prostate Cancer", *European Urology Focus*, Vol. 4/1, pp. 57-63, http://dx.doi.org/10.1016/j.euf.2016.01.016. [7]

NHS (2012), *Measuring Shared Decision Making A review of research evidence A report for the Shared Decision Making programme In partnership with Capita Group Plc Shared Decision Making*, http://www.kingsfund.org.uk/publications/nhs_ (accessed on 11 July 2019). [115]

Nordin, Å. et al. (2016), "Minimal important differences for fatigue patient reported outcome measures—a systematic review", *BMC Medical Research Methodology*, Vol. 16, p. 62, http://dx.doi.org/10.1186/S12874-016-0167-6. [58]

Norman, R. et al. (2009), "International Comparisons in Valuing EQ-5D Health States: A Review and Analysisv he_581 1194..1200", http://dx.doi.org/10.1111/j.1524-4733.2009.00581.x. [81]

Nuttall, D., D. Parkin and N. Devlin (2015), "INTER-PROVIDER COMPARISON OF PATIENT-REPORTED OUTCOMES: DEVELOPING AN ADJUSTMENT TO ACCOUNT FOR DIFFERENCES IN PATIENT CASE MIX", *Health Economics*, Vol. 24/1, pp. 41-54, http://dx.doi.org/10.1002/hec.2999. [19]

OECD (2018), *A Broken Social Elevator? How to Promote Social Mobility*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264301085-en. [1]

OECD (2018), *Health at a Glance: Europe*. [117]

OECD (2017), *Preventing Ageing Unequally*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264279087-en. [2]

OECD (2015), *Health Data Governance: Privacy, Monitoring and Research*, OECD Health Policy Studies, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264244566-en. [122]

OECD (2015), *OECD Recommendation on Health Data Governance - OECD*, https://www.oecd.org/els/health-systems/health-data-governance.htm (accessed on 18 July 2019). [20]

OECD (2014), *Geographic Variations in Health Care: What Do We Know and What Can Be Done to Improve Health System Performance?*, OECD Health Policy Studies, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264216594-en. [3]

OECD (2013), *Cancer Care: Assuring Quality to Improve Survival*, OECD Health Policy Studies, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264181052-en. [34]

OECD/EU (2018), *Health at a Glance: Europe 2018: State of Health in the EU Cycle*, OECD Publishing, Paris/European Union, Brussels, https://dx.doi.org/10.1787/health_glance_eur-2018-en. [39]

OECD/EU (2018), *Health at a Glance: Europe 2018: State of Health in the EU Cycle*, OECD Publishing, Paris/European Union, Brussels, https://dx.doi.org/10.1787/health_glance_eur-2018-en. [40]

Palmer, M. and D. Harley (2012), "Models and measurement in disability: an international review", *Health Policy and Planning*, Vol. 27/5, pp. 357-364, http://dx.doi.org/10.1093/heapol/czr047. [69]

Platt, J. et al. (2015), "Does Breast Reconstruction after Mastectomy for Breast Cancer Affect Overall Survival? Long-Term Follow-Up of a Retrospective Population-Based Cohort", *Plastic and Reconstructive Surgery*, Vol. 135/3, pp. 468e-476e, http://dx.doi.org/10.1097/prs.0000000000001054. [35]

Pusic, A. et al. (2009), "Development of a New Patient-Reported Outcome Measure for Breast Surgery: The BREAST-Q", *Plastic and Reconstructive Surgery*, Vol. 124/2, pp. 345-353, http://dx.doi.org/10.1097/prs.0b013e3181aee807. [30]

Rademakers, J. et al. (2012), "Measuring patient activation in the Netherlands: translation and validation of the American short form Patient Activation Measure (PAM13)", *BMC Public Health*, Vol. 12/1, p. 577, http://dx.doi.org/10.1186/1471-2458-12-577. [99]

Remmers, C. et al. (2009), "Is Patient Activation Associated With Future Health Outcomes and Healthcare Utilization Among Patients With Diabetes?", *Journal of Ambulatory Care Management*, Vol. 32/4, pp. 320-327, http://dx.doi.org/10.1097/JAC.0b013e3181ba6e77. [98]

Roest, A. et al. (2010), "Anxiety and Risk of Incident Coronary Heart Disease A Meta-Analysis", http://dx.doi.org/10.1016/j.jacc.2010.03.034. [72]

Rose, M. and J. Devine (2014), "Assessment of patient-reported symptoms of anxiety", *Dialogues in Clinical Neuroscience*, Vol. 16/2, p. 197, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4140513/ (accessed on 30 July 2019). [73]

Sado, M. et al. (2013), "The cost of schizophrenia in Japan", *Neuropsychiatric Disease and Treatment*, p. 787, http://dx.doi.org/10.2147/ndt.s41632. [42]

Sado, M. et al. (2013), "Cost of anxiety disorders in Japan in 2008: a prevalence-based approach", *BMC Psychiatry*, Vol. 13/1, http://dx.doi.org/10.1186/1471-244x-13-338. [43]

Schafer, W. and W. Boerma (2013), "Measures of quality, costs and equity in primary health care: instruments developed to analyse and compare primary health care in 35 countries", *Quality in Primary Care*, Vol. 21/2, http://primarycare.imedpub.com/measures-of-quality-costs-and-equity-in-primary-health-care-instruments-developed-to-analyse-and-compare-primary-health-care-in-35-countries.php?aid=186 (accessed on 14 August 2019). [124]

Scholl, I. et al. (2011), "Measurement of shared decision making - a review of instruments.", *Zeitschrift fur Evidenz, Fortbildung und Qualitat im Gesundheitswesen*, Vol. 105/4, pp. 313-24, http://dx.doi.org/10.1016/j.zefq.2011.04.012. [88]

Scurci, S. et al. (2017), "Abstract P8", *Plastic and Reconstructive Surgery - Global Open*, Vol. 5, p. 107, http://dx.doi.org/10.1097/01.gox.0000516665.93739.ee. [36]

Shaw JW, J. (2005), "US valuation of the EQ-5D health states: development and testing of the D1 valuation model", *Medical Care*, Vol. 43/3, pp. 203-20. [21]

SLOEKERS, P. et al. (2002), "A validation study of the Hospital Anxiety and Depression Scale (HADS) in different groups of Dutch subjects", *Psychological Medicine*, http://dx.doi.org/10.1017/s0033291796004382. [80]

Smagula, S. et al. (2016), *Risk factors for sleep disturbances in older adults: Evidence from prospective studies*, http://dx.doi.org/10.1016/j.smrv.2015.01.003. [62]

Stein, S. et al. (2015), "Patients' Perceptions of Care Are Associated With Quality of Hospital Care", *American Journal of Medical Quality*, Vol. 30/4, pp. 382-388, http://dx.doi.org/10.1177/1062860614530773. [8]

Sullivan, P. and V. Ghushchyan (2006), "Mapping the EQ-5D Index from the SF-12: US General Population Preferences in a Nationally Representative Sample", *Medical Decision Making*, Vol. 26/4, pp. 401-409, http://dx.doi.org/10.1177/0272989x06290496. [25]

Tan, G. et al. (2004), "Validation of the brief pain inventory for chronic nonmalignant pain", *Journal of Pain*, http://dx.doi.org/10.1016/j.jpain.2003.12.005. [53]

Targum, S., A. Nakagawa and Y. Sato (2013), "A cross-cultural comparison study of depression assessments conducted in Japan", *Annals of General Psychiatry*, http://dx.doi.org/10.1186/1744-859X-12-9. [78]

Tevis, S. et al. (2018), "Patient-Reported Outcomes for Breast Cancer", *Annals of Surgical Oncology*, Vol. 25/10, pp. 2839-2845, http://dx.doi.org/10.1245/s10434-018-6616-1. [31]

The Commonwealth Fund (2016), *International Health Policy Survey*. [45]

Thomson, S. et al. (2014), *Economic crisis, health systems and health in Europe: impact and implications for policy*, Maidenhead: Open University Press. [104]

Tirli Bryan, S. et al. (2014), *Choosing Your Partner for the PROM: A Review of Evidence on Patient-Reported Outcome Measures for Use in Primary and Community Care On behalf of the authors (see Acknowledgements)*, http://www. [46]

Trzeciak, S. et al. (2016), "Association Between Medicare Summary Star Ratings for Patient Experience and Clinical Outcomes in US Hospitals", *Journal of Patient Experience*, Vol. 3/1, pp. 6-9, http://dx.doi.org/10.1177/2374373516636681. [9]

van Hout, B. et al. (2012), "Interim Scoring for the EQ-5D-5L: Mapping the EQ-5D-5L to EQ-5D-3L Value Sets", *Value in Health*, Vol. 15/5, pp. 708-715, http://dx.doi.org/10.1016/j.jval.2012.02.008. [24]

Van Reenen, M. and B. Janssen (2015), *EQ-5D-5L User Guide Basic information on how to use the EQ-5D-5L instrument*. [71]

Veroff, D., A. Marr and D. Wennberg (2013), "Enhanced Support For Shared Decision Making Reduced Costs Of Care For Patients With Preference-Sensitive Conditions", *Health Affairs*, Vol. 32/2, pp. 285-293, http://dx.doi.org/10.1377/hlthaff.2011.0941. [12]

Wang, K. et al. (2016), "Research on translations of tests: Cross-cultural validation of the depression anxiety stress scale-21 in China", *Psychological Assessment*, http://dx.doi.org/10.1037/pas0000207. [79]

Ward, B. and L. Black (2016), "State and Regional Prevalence of Diagnosed Multiple Chronic Conditions Among Adults Aged ≥18 Years — United States, 2014", *MMWR Morb Mortal Wkly Rep*, Vol. 65, pp. 735–738. [101]

Ware, J. and B. Gandek (1998), "Overview of the SF-36 Health Survey and the International Quality of Life Assessment (IQOLA) Project", *Journal of Clinical Epidemiology*, Vol. 51/11, pp. 903-912, http://dx.doi.org/10.1016/S0895-4356(98)00081-X. [82]

Watkins, L. (2014), *Aligning payers and practices to transform primary care: a report from the Multi-State Collaborative*, New York (NY) : Milbank Memorial Fund. [105]

Weiland, T. et al. (2015), "Clinically significant fatigue: prevalence and associated factors in an international sample of adults with multiple sclerosis recruited via the internet.", *PloS one*, Vol. 10/2, p. e0115541, http://dx.doi.org/10.1371/journal.pone.0115541. [61]

Whitehead, L. (2009), "The Measurement of Fatigue in Chronic Illness: A Systematic Review of Unidimensional and Multidimensional Fatigue Measures", *Journal of Pain and Symptom Management*, Vol. 37/1, pp. 107-128, http://dx.doi.org/10.1016/J.JPAINSYMMAN.2007.08.019. [59]

Wong, E. et al. (2019), "Associations between provider communication and personal recovery outcomes", *BMC Psychiatry*, Vol. 19/1, p. 102, http://dx.doi.org/10.1186/s12888-019-2084-9. [11]

Wong, S. et al. (2014), "Patient-reported confidence in primary healthcare: are there disparities by ethnicity or language?", *BMJ Open*, Vol. 4, p. 3884, http://dx.doi.org/10.1136/bmjopen-2013. [116]

Zaslansky, R., W. Meissner and C. Chapman (2018), "Pain after orthopaedic surgery: differences in patient reported outcomes in the United States vs internationally. An observational study from the PAIN OUT dataset", *British Journal of Anaesthesia*, Vol. 120/4, pp. 790-797, http://dx.doi.org/10.1016/j.bja.2017.11.109. [55]

Zhuo, L. et al. (2018), "Time Trade-Off Value Set for EQ-5D-3L Based on a Nationally Representative Chinese Population Survey", *Value in Health*, Vol. 21/11, pp. 1330-1337, http://dx.doi.org/10.1016/j.jval.2018.04.1370. [22]

Zoëga, S., S. Ward and S. Gunnarsdottir (2014), "Evaluating the Quality of Pain Management in a Hospital Setting: Testing the Psychometric Properties of the Icelandic Version of the Revised American Pain Society Patient Outcome Questionnaire", *Pain Management Nursing*, Vol. 15/1, pp. 143-155, http://dx.doi.org/10.1016/j.pmn.2012.08.001. [56]

[1] Based on 45,600 total hip replacements reported in 2016 and 49,500 TKR in 2017 at a 'national efficient price' – (the official price governments for conducting these procedures in the public sector) of just under AUD 20,000 for each procedure (IHPA 2019). The national figure is likely to be higher because approximately half of procedures are carried out in the private sector where higher prices are typically paid.

[2] Coxa hospital has a patient catchment covering an entire region of Finland.

[3] With the exception of Galeazzi, which included all principal diagnoses.

[4] The value is derived by subtracting the pre-operative score from the post-operative score. A positive value therefore represents an improvement in QoL.

[5] The degree of improvement was statistically meaningful at the 95% confidence level in all programmes and in aggregate.

[6] The generic and condition-specific scales are not linear – i.e. a change from 0.2 to 0.3 is not necessarily the same magnitude in terms of health-related QoL than 0.7 to 0.8. The percentage improvements are provided for illustrative purposes and should be interpreted cautiously.

[7] HOOS-PS: Hip disability and Osteoarthritis Outcome Score–Physical Function Shortform

[8] An alternative scoring system exists for both instruments where a *lower* value represents a *better* result.

[9] See 6.

[10] KOOS-PS: Knee injury and Osteoarthritis Outcome Score-Physical Function Shortform

[11] See 6.

[12] See 6.

[13] As valued by a US population sample (Shaw JW, 2005).

[14] Based on the average life expectancy at age 65 in the countries of the contributing programs of 20.5 years, subtracting one year to account for recovery and rehabilitation.