Synthesis of the literature on patient-reported outcomes (2010-2019)

Madeleine Akrich
CSI - Centre de sociologie de l’innovation, i3 UMR CNRS
Mines ParisTech, PSL Research University
madeleine.akrich(a)mines-paristech.fr

Florence Paterson
CSI - Centre de sociologie de l’innovation, i3 UMR CNRS
Mines ParisTech, PSL Research University
florence.paterson(a)mines-paristech.fr

Vololona Rabeharisoa
CSI - Centre de sociologie de l’innovation, i3 UMR CNRS
Mines ParisTech, PSL Research University
vololona.rabeharisoa(a)mines-paristech.fr

Working Paper 20-CSI-01_EN
April 2020

L’institut interdisciplinaire de l’innovation

(UMR 9217) a été créé en 2012. Il rassemble :

- les équipes de recherche de MINES ParisTech en économie (CERNA), gestion (CGS) et sociologie (CSI),
- celles du Département Sciences Economiques et Sociales (DSES) de Télécom ParisTech,
- ainsi que le Centre de recherche en gestion (CRG) de l’École polytechnique,

soit plus de 200 personnes dont une soixantaine d’enseignants chercheurs permanents.

L’institut développe une recherche de haut niveau conciliant excellence académique et pertinence pour les utilisateurs de recherche. Par ses activités de recherche et de formation, i3 participe à relever les grands défis de l’heure : la diffusion des technologies de l’information, la santé, l’innovation, l’énergie et le développement durable. Ces activités s’organisent autour de quatre axes :

- Transformations de l’entreprise innovante
- Théories et modèles de la conception
- Régulations de l’innovation
- Usages, participation et démocratisation de l’innovation

Pour plus d’information : http://www.i-3.fr/

Ce document de travail est destiné à stimuler la discussion au sein de la communauté scientifique et avec les utilisateurs de la recherche. Son contenu est susceptible d’avoir été soumis pour publication dans une revue académique. Il a été examiné par au moins un referee interne avant d’être publié. Les considérations exprimées dans ce document sont celles de leurs auteurs et ne sont pas forcément partagées par leurs institutions de rattachement ou les organismes qui ont financé la recherche.

The Interdisciplinary Institute of Innovation

(UMR 9217) was founded in 2012. It brings together:

- the MINES ParisTech economics, management, and sociology research teams (from CERNA, CGS and CSI),
- the Department of Economics and Social Science (DSES) at Télécom ParisTech, and
- the Management Research Center (CRG) at Ecole Polytechnique.

Collectively, more than 200 people are affiliated with these entities, including 60 permanent academic researchers.

i3 performs high-level research, combining academic excellence and relevance for end users. Through its teaching and research activities, i3 actively seeks to address current challenges in several areas: the diffusion of communication technologies, health, innovation, energy, and sustainable development. These activities are organized around four main topics:

- The transformation of innovating firms,
- Theories and models of design,
- The regulation of innovation, and
- The use of, participation in, and democratization of innovation.

For more information: http://www.i-3.fr/

This working paper is intended to stimulate discussion within the research community and among research users. Its content may have been submitted for publication in academic journals. It has been reviewed by at least one internal referee before publication. The views expressed in this paper are those of the author(s) and not necessarily those of the host institutions or funders.
ABSTRACT

The aim of this synthesized literature review is to provide an overview of patient-reported outcomes (PROs), a term we use to refer to both health outcomes reported directly by patients, and the tools used for their collection and measurement. Developed from the 2000s onwards, PROs have multiplied as their uses have diversified. We review approximately 50 articles authored by clinical researchers, public health physicians, and to a lesser extent, health economists, as well as approximately 10 practice guidelines distributed by public agencies, research consortia, and think tanks that develop PROs. The overview begins with descriptions of a few historical milestones that reveal the decisive role played by the FDA in the use of PROs when reviewing new treatments tested in clinical trials. These descriptions are followed by a formal definition of PROs that has now been adopted by most of the parties involved. The review continues with a focus on the different categories of PROs developed over the last two decades, well beyond the framework of clinical trials alone. Finally, it addresses recent debates on the capacity of PROs to collect and measure patient experiences, the methodological issues involved in the design of these tools, and in particular, the participation of patient organizations in the construction of PROs that better reflect the specific concerns of the people affected.

KEY-WORDS

Patient-reported outcomes (PROs), health policies, clinical trials, methodology, patient experience measurement, patient organizations.
Synthesis of the literature on patient-reported outcomes (2010-2019)*

Introduction

The aim of this synthesized literature review is to provide an overview of patient-reported outcomes (PROs), a term we use to refer to both health outcomes reported directly by patients, and the tools used for their collection and measurement. Since their institutionalization in the 2000s, PROs have multiplied, as have the public agencies, research consortia, and think tanks that develop PROs and publish guidelines for best practices. Patient organizations are also involved in the production and use of PROs, often in collaboration with expert groups. Over the last two decades, a constellation of instruments has been developed, each of which has been assigned its own label and designed for a particular use. The aim of this synthesis is to provide an overview of these instruments and related discussions.

The literature on PROs has been authored primarily by clinical researchers, public health physicians, and to a lesser extent, health economists. This synthesis is based on a selection of approximately 50 articles, as well as approximately 10 practice guidelines and working group reports. This synthesis is composed of three parts. The first part describes a series of historical milestones in the development of PROs, highlighting the decisive role played by the FDA in proposing the use of PROs in the clinical assessment of new treatments tested in clinical trials and in establishing a definition of PROs that has now been adopted by most of the parties involved. The second part describes the different categories of PROs that have been developed over the last two decades and which today extend beyond the scope of clinical trials alone. The third part examines more recent debates on the capacity of PROs to collect and measure the actual experiences of patients, and the involvement of patient organizations in the construction of PROs that better reflect the concerns of the people affected. The main findings are summarized at the end of each section or sub-section. The corpus of literature on which this synthesis is based is included at the end of this document.

* This synthesis has been prepared by following the framework of the TRT_cSVD Project “From Target Identification to Next Generation Therapies for Small Vessel Disease”– WP7 Ethical and Sociological Approaches to CADASIL and other familial SVDs. Convention N° 16-RHUS-0004.
I/ Historical milestones

PROs originated from two sets of initiatives aimed at (i) measuring the effects of health policies, and (ii) evaluating the effectiveness of new drugs in clinical trials and pharmacovigilance.

I.1/ Assessing the effects of health policies on patients’ quality of life

Benamouzig (2010) provided historical background on the measurement of the effects of health policies. In France, as in other countries, these measurement tools were initially developed as part of post-war national planning efforts and privileged the views of scientific experts. From the 1970s onwards, the views of individuals were gradually integrated, a phenomenon that was accentuated in the 1980s as health expenditures were increasingly rationalized. This consideration of individuals’ points of view was based on the observation that, in addition to the possible clinical effects of health policies, it was necessary to assess the impacts of these policies on people’s health-related quality of life (HRQoL), a criterion now considered important for the effective allocation of resources. This approach was advocated by RAND, an expert organization that advises the U.S. federal government, in a study commissioned in 1971 on the impacts of alternative forms of health insurance schemes on health status. In the United Kingdom, economist Alan Williams and psychiatrist Rachel Rosser also developed the famous quality-adjusted life year (QALY) indicator to assess therapeutic innovations with a view to their possible funding by the NHS. The British National Institute for Clinical Excellence (NICE), which was established in 1999 and is responsible for what is known as Health Technology Assessment, bases its recommendations on this indicator of the utility of new treatments by incorporating the subjective preferences of individuals. France has not adopted this utilitarian approach and favors a strictly medical assessment.

These initiatives highlight two components of the historical landscape within which PROs have developed. The first component addresses the issue of the value and usefulness of health policies, or more narrowly, of new treatments for patients. This question has been the subject of many debates and has led different consortia to propose clarifications. For example, the International Consortium for Health Outcomes Measurement (ICHOM) proposed the following definition of “value-based healthcare” in 2018:

Value is defined as the outcomes that patients experience relative to the cost of delivering those outcomes. Value-based health care, or VBHC, is health care that delivers the best possible outcomes to patients for the lowest possible cost. We believe that choice and competition in health care should be based on value. By restructuring care-delivery around outcomes, and promoting superior outcomes with financial incentives, health care systems will improve quality and curb inefficiencies. This will benefit every stakeholder across the health care spectrum. In a value-based world:

- patients are able to choose providers based on informed expectations of outcomes and the associated costs;
- providers that deliver superior outcomes at competitive costs thrive, while others improve or lose market share;
- payers negotiate contracts based on results and encourage innovation to achieve those results;
suppliers succeed by marketing their products on value, showing improved patient outcomes relative to costs. (ICHOM, 2018. What is value-based health care? Available on the ICHOM’s website at https://www.ichom.org/faqs/)

This definition echoes that of utility proposed by the European Patient’s Academy on Therapeutic Innovation (EUPATI) in 2016:

Utility is the (perceived) capacity of an item to satisfy needs or wants. In health economics, utilities measure the intensity of patient preferences. For example, the importance of various factors to patients, such as symptoms, pain, and psychological health. The impact of new treatments on these factors, and therefore on quality of life (QoL), can then be calculated. This is a common approach used by health technology assessment (HTA) agencies that advise whether or not treatments should be funded (for example) by government health services. (EUPATI, 2016)

The notions of “value-based healthcare” and utility are therefore based on a vision of healthcare strategies as part of a market in which patients behave like economic agents who have preferences, are able to articulate their needs, and consequently are able to identify therapies that best meet their expectations and advise funding authorities accordingly.

The second component of these initiatives to measure the effects of health policies is that they are based on questionnaire surveys of patients. Two generic questionnaires, as opposed to pathology-specific questionnaires, constitute references for the community of experts and decision-makers.

The first questionnaire is the EQ-5D, which was developed between 1987 and 1991 by EuroQol Group, a multidisciplinary and international group of researchers (EuroQol Group, 1990). Patients who complete the EQ-5D rate five different dimensions of their health—mobility, self-care, usual activities, pain and discomfort, anxiety, and depression—on a scale of 1 to 5 (Desomer et al., 2018), and rate their overall health status on a scale of 0 (worst possible health) to 100 (best possible health). This type of generic tool contributes to cost-benefit analyses (Devlin & Brooks, 2017) and is widely used by NICE, the UK’s health technology assessment agency, to inform decision-making about health products and technologies (Haywood et al., 2016).

The second questionnaire is the SF-36 developed based on the RAND Medical Outcomes Study (1987–1989). The SF-36 is commonly used in health-related quality of life studies. RAND Health Care provides three versions of a free questionnaire (12, 20, or 36 items) covering, for example, physical functioning (measuring physical limitations and capacities, mobility), compliance, sleep, sexuality, mental health, etc. (Gregory et al., 2018). The SF-36 yields separate scores for physical and mental health, rather than an overall health score (Devlin et al., 2010). The instrument has been translated into many languages and has given rise to many adaptations.

These two questionnaires have two characteristics in common: (i) they ask patients to evaluate standardized items selected by medical experts to measure health-related quality of life; and (ii) they translate evaluations into quantitative indicators that make it possible to compare patients’ assessments of different care strategies. Patients are thus considered to be producers of data considered relevant by experts.
Linking PROs with measurements of the effects of health policies is part of a movement that can be described as "evidence-based policy," which is motivated by a concern for better allocation of resources and based on patients’ reports of the quality of life provided (or not) by different care strategies.

I.2/ Assessment of the therapeutic benefits of treatments tested in clinical trials

The second set of initiatives from which PROs have developed relates to the assessment of the efficacy of new drugs in clinical trials and pharmacovigilance. These initiatives introduce patient reports on the benefits of treatments into such evaluations, both pre- and post-marketing.

In the United States, RAND played an important role in the development of a “science of health outcomes measurement” in the 1980s. Pharmacological researchers also began to consider quality of life as a relevant outcome to be measured in the assessment of new drugs around the same time. Likewise, the FDA was asking sponsors of new oncology products to consider quality of life measures formulated in terms of symptoms and functioning based on the idea that traditional objective measures, such as tumor response, do not always reflect the benefits perceived by patients. In 1985, a medical journal published the results of a clinical trial in which quality of life was used as the primary endpoint for the first time; specifically, the researchers measured the “tolerability” of the product rather than its efficacy (Johnson & Temple, 1985). In response to the frenzy that this publication generated in the industry, the FDA established a working group and consulted with interested professionals to formulate principles to inform the use of PROs and provide a framework for product evaluation. This led to the development of guidelines which have been amended over the years. In 2000, these guidelines defined a PRO, thereby distinguishing it from a quality of life measurement. This definition was updated in 2009 and adopted by the European Medicines Agency (EMA) and the academic world:

A PRO is any report of the status of a patient’s health condition that comes directly from the patient, without interpretation of the patient’s response by a clinician or anyone else. The outcome can be measured in absolute terms (e.g., severity of a symptom, sign, or state of a disease) or as a change from a previous measure. In clinical trials, a PRO instrument can be used to measure the effect of a medical intervention on one or more concepts (i.e., the thing being measured, such as a symptom or group of symptoms, effects on a particular function or group of functions, or a group of symptoms or functions shown to measure the severity of a health condition). (FDA, 2009)
This definition indicates the key characteristics of PROs: the concept refers to significant therapeutic benefits to be measured in clinical trials; the symptom refers to “subjective evidence of a disease, condition or treatment-related effect that can only be perceived and known by the patient;” the sign refers to “objective evidence of a disease, condition or treatment-related effect,” “usually observed and interpreted by the clinician,” but which the patient can “notice” and “report.”

The notion of PRO was introduced in the 2000s, as part of the assessment of experimental treatments tested in clinical trials. In addition to measures of quality of life, PROs include elements related to manifestations of a disease and the effects of a treatment as perceived by patients. Since then, PROs have gained momentum and have become the subject of extensive discussion.

II/ The multiplication of PROs and their uses

The development of PROs from the 2000s onwards has manifested in several ways. First, many institutional actors, researchers in biomedicine, economics, and health management, think tanks, and patient organizations have produced PROs, organized themselves into consortia and working groups to discuss them, and published practice recommendations. Second, as PROs have multiplied, new uses beyond the scope of clinical trials have emerged, and researchers have explored their application to a broad and diverse set of other contexts. Reflecting on the design of PROs, the FasterCures think tank notably proposed a workshop in 2017 entitled “Patient-Reported Outcomes: Design with the End in Mind” (Grossman, 2018). Based on the literature, we can distinguish four ways of categorizing PROs and their uses, and the specific questions that each type of PRO raises.
II.1/ Categorizing PROs according to the levels at which they are used

A/ The macroeconomic or system level

PROs can be used to monitor the health of the population, to design and evaluate public health policies (prevention, health promotion, measurement of health disparities, etc.), and to inform decisions about the reimbursement of care-related expenses, the contracting of health services and pay-for-performance. According to Desomer et al. (2018), what increases the value of PROs is the possibility of linking them to monitoring data in clinical registries, or data related to health care expenses and length of hospital stays. Such links make it possible to cross-reference different medical and economic data and to study the correlations between, for example, clinical data collected by health care professionals and compiled in registries, and PROs.

B/ The institutional or service level

PROs can be used to pilot quality improvement initiatives. Aggregate data can be used to evaluate and compare provider performance (benchmarking and feedback), identify quality issues that remain under-treated, and inform the general public with the aim of promoting informed patient choice (Desomer et al., 2018). This type of use has developed mainly in the United Kingdom. However, critics have argued that this type of analysis fails to clearly demonstrate links between PROs and quality of care. Extensive effort has been directed toward measuring quality of care, mainly in terms of assessing health care services according to criteria such as safety, responsiveness, or efficiency. This type of criticism illustrates how difficult it is to convince experts of the validity of non-traditional criteria for both assessing the quality of health care services and informing the organization of such services.

C/ Clinical practice level or individual level

PROs can be used to refine a diagnosis (Desomer et al., 2018), as a support for shared decision-making between caregiver and patient (McKenna, 2011), to engage the patient in health management (Snyder et al., 2012), and to report problems that would not otherwise be addressed, such as anxiety and depression, and problems with physical, social, or emotional functioning (Aaronson et al., 2011; Greenhalgh, 2008). PROs also can be used to facilitate patient-centered discussions in multidisciplinary team meetings (Aaronson et al., 2011; Desomer et al., 2018; Greenhalgh, 2008; Snyder et al., 2012). Notably, although PROs allow for the identification and discussion of problems encountered by patients during consultations, they have little impact on how clinicians subsequently manage these problems.¹

II.2/ Categorizing PROs according to their intended objectives

A/ Measuring the effects of an experimental treatment in a clinical trial

When clinical trials involve treatments for which there are no objective outcome measurements (degree of morbidity, biomarkers of symptoms, etc.) and for which results can only be

¹ The International Society for Quality of Life Research publishes a “User’s Guide to Implementing Patient-Reported Outcomes Assessment in Clinical Practice” (Aaronson et al., 2011).
observed subjectively in terms of patient impacts, PROs can be used as primary outcome measures. PROs are also used to complement primary outcomes (e.g., survival rates, biomarkers), as they reflect outcomes that are important to patients and may reveal relationships between symptoms and other indicators such as quality of life (Weldring & Smith, 2013).

**B/ Identifying treatment side effects**

**Patient-reported outcomes of adverse events (PRO-AEs)** are used to identify treatment side effects. The Patient-Reported Outcomes Safety Event Reporting (PROSPER) Consortium was established to improve drug safety by adopting an approach that fully integrates the patient’s perspective according to a “patient-centered model of care.” The consortium has developed recommendations and guidelines to promote the development of PRO-AEs, and more broadly, devices for capturing and analyzing patient-reported data. According to the consortium, the benefit of patient reporting is that patients respond earlier than health professionals, perceive the impacts and severity of drug reactions differently, and provide more detailed descriptions of the circumstances surrounding drug use, including cases of prescribing errors or unintended use. The aim is to obtain "real-world" data to improve the assessment of the benefits and risks of pharmaceutical therapies (Banerjee et al., 2013).

**C/ Early detection of alerts reported by patients regarding their health status**

Research has been undertaken to assess health care service responses to **PRO-alerts**, urgent updates provided by patients regarding their health status. One of the conclusions of this research is that research nurses, data managers/coordinators, trial managers, or chief/principal investigators who become aware of these alerts when processing PRO data are very often uncertain about how to respond. Moreover, these alerts generally are not processed in real time. The results of the STAR study, presented by oncologist Ethan Basch at the annual conference of the American Society for Clinical Oncology in 2017, have generated interest among members of the medical press (Dean et al., 2017; Marino et al., 2018). The study reveals the effects of routine electronic reporting by cancer patients on their health status and health-related quality of life. The electronic reporting system was able to notify a nurse who would contact the patient, answer questions, and suggest a solution or consult with a team physician if necessary.

**D/ Collecting patient-reported data for clinical research**

The fourth objective of collecting patient-reported data for clinical research is greatly facilitated by information and communication technologies and is part of efforts to harness the power of "big data" in the field of health.

**Patient-reported information (PRI)** is presented as an alternative to PROs. Unlike traditional PROs, the novelty lies in the fact that information can be collected and disseminated using social networking platforms (Baldwin et al., 2011).

**Patient-generated health data (PGHD)** are recorded and transmitted by patients to teams of researcher-clinicians using smart phones, digital platforms, and other technologies. PGHD have five characteristics: (i) patients (or trusted third parties, if the patients cannot do so) record
and transmit these data, not caregivers; (ii) data are collected outside clinical settings; (iii) data are longitudinal, and it is possible to obtain repeated measurements over time; (iv) the frequency of collection may be very high; and (v) depending on the parameter of interest, the data flow can be quasi-continuous over long observation periods (Wood et al., 2015).

II.3/ Categorizing PROs according to their degree of generality

The literature distinguishes between two types of PROs: generic PROs and specific PROs.

A/ Generic PROs

Generic PROs are assessed using a specific set of questions, regardless of patients‘ health status. The aim is that health status (and changes in health status, in cases where generic PROs are measured before and after treatment or several times during treatment) can be compared between patients or population groups, or between care sectors. Generic PROs measure the impacts of health status on life in general, quality of life, and quality of life relative to certain dimensions such as physical functioning. The lack of robustness of comparisons conducted on the basis of generic PROs, as well as the obsolete nature of certain items (which sometimes date back several decades) are highlighted (McKenna, 2011). To address these shortcomings, an item-banking approach can be adopted whereby parts of questionnaires designed to assess generic PROs are replaced with domain-relevant items (McKenna, 2011).

B/ Condition-specific or disease-specific PROs

Condition-specific or disease-specific PROs relate to a particular disease (diabetes), group of diseases (cancer), body part (the knee), function (visual, auditory, etc.), or focus of care (pain). Questions relate to the severity of the condition and are therefore tailored to the people affected. For this reason, they are not used in population health surveys. Items are designed based on qualitative interviews with affected patients and tested with new patient populations to assess their validity. A well-designed measurement tool only includes questions that are relevant, meaningful and acceptable to patients. Research shows that disease-specific PROs are more likely to show differences between competing therapies (Desomer et al., 2018; Devlin et al., 2010).

The different types of PROs and their uses, which have multiplied since the 2000s, show: (i) the progressive institutionalization of PROs at different levels of decision-making (i.e., the health system, health services, medical practices), and for different activities (i.e., clinical trials, pharmacovigilance, patient feedback on their health status, clinical research); (ii) a variety of methods for collecting PROs and the nature of the data collected (e.g., standardized questionnaires with items defined by experts, data collected through digital self-tracking equipment, qualitative interviews with patients, information disseminated by patients via social networks); and (iii) necessary adjustments to practices and the organization of care to ensure PROs are adequately considered (e.g., the organization of responses to PRO-alerts). Beyond the specific concerns raised by each PRO, a cross-cutting question has emerged in recent years: To what extent do PROs accurately reflect patients‘ experiences with diseases or treatments?
III/ From PROs to patients’ experiences

Although PROs are being designed and applied in an increasing number of contexts, experts have expressed reservations since their inception. Specifically, PROs reveal patients’ subjective perceptions and do not constitute reliable data for objectively assessing health conditions and the effects of health policies or treatments on health status. For this reason, a number of consortia have been established to develop PRO design methodologies that neutralize these subjective biases by offering standardized and validated questionnaires and graduated scales on which patients can objectively assess their perceptions. McKenna (2011), a Research Director at an organization specializing in PROs (Galen Research Ltd., Manchester), emphasized that when developing PROs, “common sense” is insufficient; scientific expertise is required, as PROs must be designed in a way that enables qualitative information to be quantified. This is also why certain experts consider PGHD, which are collected and transmitted by patients via digital self-tracking equipment, to be much more reliable than traditional PROs based on subjective patient reports. This question of the subjectivity of PROs is the focus of ongoing discussions and reflections by organizations and actors involved in their production and use (Shapiro et al., 2012).

Mirroring the issue of the subjectivity of patients’ reports, a second topic of discussion has emerged in recent years regarding the relevance of the questions patients are asked. Such discussions have been motivated by the development of “patient-centered healthcare” and by the increased participation of patients in health issues, which is seen as both an ethical and democratic imperative. Concerns over the relevance of questions point to possible discrepancies between the PRO items developed by medical experts and the experiences of patients, which may not necessarily correspond to these items. This has led to the development of a set of tools known as patient-reported experience measurements (PREMs), and of principles and recommendations for the design and use of these tools. It has also led to the involvement of patient organizations in the construction of such tools.

III.1/ How can patients’ experiences be taken into account?

A number of experts have noted a discrepancy between the questions patients are asked and their experiences of a disease or treatment. For example, together with his colleagues, Dr. Ethan Basch, Director of the Cancer Outcomes Research Program and Professor of Medicine and Public Health at the UNC Lineberger Comprehensive Cancer Center in North Carolina, found that the side effects observed during phase 1 of a clinical trial measuring the toxicity of an experimental treatment were not included in the Common Terminology Criteria for Adverse Events (CTCAE) of the National Cancer Institute (NCI). As they began to record symptoms in patients, they discovered that patients were detecting more symptoms than clinicians, and doing so much earlier (Basch et al., 2015). Basch, Abernethy, and Reeve (2011) discussed the example of a pain item that was originally designed for use by clinicians and was included in a PRO without undergoing a qualitative assessment involving direct patient participation. The terms on the pain scale below are problematic because they merge different registers—that is, the words typically are not used together to draw comparisons (e.g., mild, discomforting, and horrible). Moreover, the distinctions between, for example, mild and
discomforting or horrible and excruciating are not explicit. They concluded that the ability of this item to assess patients’ experiences of pain associated with both disease and treatment is compromised.

<table>
<thead>
<tr>
<th>How strong is your pain?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mild</td>
</tr>
</tbody>
</table>

Present Pain Intensity (PPI) item of the McGill Pain Questionnaire

In light of these findings, the solution recommended by several experts and regulatory authorities such as the FDA and certain PRO research and production consortia is to encourage direct patient participation in the development of the items. For example, the FDA’s “Patient Focused Drug Development Efforts” program publishes guidelines for collecting data about patient experiences. The agency also offers workshops to review what it calls **clinical outcome assessments (COA)** which include PROs and **real-world evidence (RWE)** (Treadwell, 2018). The PROSPER Consortium has published a report indicating that PROs should include not only patient reports collected by clinicians using validated instruments, but also the full range of voluntary patient reports.

For now, it seems as though these recommendations are being disregarded. Wiering, de Boer, and Delnoij (2017) investigated patient participation in the development of PROs by examining 189 studies. Patient participation was not documented a quarter of the studies, particularly older ones. When it was documented, patients were involved mainly in question development (58%) and comprehension testing (51%). In contrast, patient involvement in identifying outcomes to be measured (“concepts of interest,” in the FDA’s terminology) was very low (11%). Patients were involved in all aspects of PRO development in only 7% of cases (Wiering et al., 2017). Even more instructive, when standards for the use of PROs in randomized controlled trials were set by the International Society for Quality of Life Research (ISOQOL), the issue of patient involvement was rejected during the review process. Dean et al. (2017) suggested that change could only come from learning how to listen to patients, both as representatives on the various committees and as researchers, policy makers, and industrialists.

With regard to learning how to listen to patients, the question arises of how to acquire a better understanding of their experiences. Some authors have called for reflection on the use of qualitative research methodologies applied in sociology and anthropology (i.e., in-depth interviews, focus groups, observations, use of visual aids) to be implemented in the initial phases of studies (Dean et al., 2017). Others have suggested interviewing patients about specific experiences rather than assessing their satisfaction (Nilsson et al., 2016). Still others have suggested that a distinction should be made between classic PROs and what they call **patient-centered PROs (PC-PROs)**: patients report information in both cases, but the difference lies in what information is reported (Tractenberg et al., 2017). A PC-PRO, the authors

---

2 The FDA defines real world data (RWD) as care information that comes from multiple sources generally outside the typical clinical research setting, such as electronic medical records, complaint and billing data, product and disease registries, or data collected through personal equipment and health-related applications.
explained, “explicitly maintains the primacy of the patient’s, vs. the clinician’s or researcher’s, perspective on the patient’s health status, condition, or experience” (Tractenberg et al. 2017, p. 3). They detailed the iterative process they implemented to build PC-PROs in partnership with interested parties (i.e., experts representing the target population, caregivers, clinicians, and researchers). The process begins by compiling an initial list of signs and symptoms based on a literature review and discussions with patients, researchers, and clinicians; this list, which is amended throughout the process, is designed as a working tool. Focus groups are organized with both patients and clinicians. Particular attention is paid to maintain alignment with the perspectives of patients and clinicians on what to measure and how to measure it independently, with the ultimate goal of integrating the patient experience as captured by all participants in the design of PC-PROs (Tractenberg et al., 2017).

PROs have been designed to collect patients’ responses to questions developed by experts to augment knowledge of the effects of treatments or health policies on people’s lives. Considerable effort has been made to standardize questions and validate the answers in order to neutralize the subjective nature of patient reports. Recently, there has been important reflection not so much on the quality of the answers, but on the relevance of the questions, and the idea is beginning to spread that PROs should be designed to collect patients’ experience of a disease or treatment without preconceived orientations. Initiatives have been taken in this direction, contributing to the emergence of PREMs.

III.2/ The role of patient organizations in the development of PROs

As the literature demonstrates, experts are advocating that patients should be involved in the development of PROs. From this perspective, patient associations are well-positioned to facilitate this process, given their role in collecting and shaping the experiences of patients. Some experts have recognized this unique capacity of patient associations. For example, Nicklin and his colleagues at the University of the West of England, Bristol, have developed an approach to capture the fatigue experience of rheumatoid arthritis patients:

Patients have an understanding of fatigue and its effect on their everyday life that only they can describe. Therefore, an understanding of the lived experience of rheumatoid arthritis (RA) fatigue could be gained through listening to patients (experts) and applied to the development of PROMs [patient-reported outcome measures], which will then capture issues relevant to patients. (Nicklin et al., 2010, p. 1553).

According to these researchers, not only do commonly-used tools use different vocabularies, which makes comparisons difficult, but more importantly, they only measure the severity of symptoms (e.g., fatigue) and not the perceived impacts of these symptoms on the ability to cope, which may be equally (if not more) important to patients. They have established a partnership between researchers and an organization of patients affected with RA, and a list of terms used by patients to describe fatigue during qualitative interviews has been gradually established and refined.
Morel and Cano (2017), members of the IRDiRC (International Rare Diseases Research Consortium), considered it essential to produce knowledge on rare diseases, which are often poorly known or incompletely described, or described in ways that do not accurately reflect patients’ experiences. In addition to qualitative interviews and focus groups, they recommended the use of patient diaries to design PROs that are more relevant to patients:

It is important to invest in a careful description of the clinical manifestations, disease course, clinical outcomes and – importantly – of the disease impact on patients’ daily life and of the patients’ chief complaints and expectations from future therapies. A comprehensive understanding of patient disease burden is key to later support the assessment and selection of the outcome measures that are most relevant to patients. Because a rare disease affects every aspect of their daily life, patients and their caregivers become experts of the rare condition and of the important outcomes of diseases that need to be addressed. It is thus critically important to partner with and listen to them. (Morel & Cano, 2017, p. 3)

In order to effectively incorporate patients’ experiences into PROs, it is recommended that core outcome sets (COSs) be identified, i.e., the essential outcomes that patients expect from new drugs or care strategies (Dean et al., 2017; Mills et al., 2016).

Patient organizations are launching initiatives as well. For example, the Michael J. Fox Foundation for Parkinson’s Research collected data directly from patients and caregivers about their daily experiences with Parkinson’s disease using Apple’s ResearchKit and a digital platform (Grossman, 2018). Initiatives are multiplying, such as the Duchenne Parent Project’s outcome measurement initiative, the Patient Reported Outcomes Burdens and Experiences on Hemophilia, the Dravet syndrome internet platform, and the survey on the perspectives and priorities of patients with spinal muscular atrophy conducted by SMA Europe and Cure SMA (Morel & Cano, 2017).

PROs in the field of rare diseases deserve a special mention. In the context of orphan drug development, the EMA requires clinical trial documentation to include a pre-specified PRO endpoint in addition to other clinical efficacy endpoints, and suggests the use of Health-Related Quality of Life (HRQoL) questionnaires. An examination of the Summaries of Product Characteristics (SmPCs) of orphan drugs approved by the EMA from 2012 to 2016 revealed that only 27% of these documents mention the use of PROs, compared to more than 45% for documents relating to non-orphan drugs (Jaroslawski et al., 2018). The same trend can be observed for FDA-approved orphan drugs. The authors offered an explanation for this phenomenon: in order to speed up the availability of orphan drugs, less evidence is provided for marketing authorization; therefore, pharmaceutical companies look for less data during the product development process, including data that PROs can provide. Benjamin and colleagues (2017) provided a different explanation: given the rarity of the diseases, the fact that many of them affect young children, and the wide heterogeneity of clinical situations, specific PROs must be developed for each disease or group of diseases (e.g., rare metabolic diseases), which are currently very few in number. In their report on how to proceed with the development of PROs in the field of rare diseases, they advocated:

• The introduction of observer-reported outcomes (ObsROs) which a third party could provide if the patient is unable to do so, either because the individual is too severely affected or because the individual is a child;
• The mobilization of patients’ associations during the upstream phases of the construction of PROs;
• The development of PROs that would provide information on three sets of phenomena:
  o The natural history of diseases, based on the collection of patients’ and families’ experiences, as well as data from disease registries (some of which are set up by patient associations), literature reviews, satisfaction surveys among the people concerned, etc.;
  o The therapeutic benefits of treatments from patients' perspectives, focusing on the common benefits sought by the population of people concerned, especially those sought in the short term. Patient-reported benefits could then be used as “surrogate outcomes” in the evaluation of drug efficacy in order to overcome the difficulties of conducting randomized double-blind clinical trials in the case of rare diseases.
  o The construction of specific aggregation methods for the data collected and/or the adaptation of methods used in generic PROs such as SF-36 or EQ-5D, in order to account for data heterogeneity. For example, the EuroQoL group has just developed a child-friendly version of the EQ-5D. Likewise, Basch and Bennett (2014) described a multi-attribute questionnaire whereby a numerical value is used to weight the most sought-after attributes in the subject studied. Such a questionnaire was designed to aggregate data collected on a drug called Ruxolitinib which was approved in 2011 by the FDA and EMA for patients with myelofibrosis.

Finally, it should be noted that in the field of rare diseases, patient organizations have launched initiatives to self-test medications, the results of which have been measured using patient reports. The best-known case is the amyotrophic lateral sclerosis (ALS) lithium study conducted on the online platform PatientsLikeMe (Stefanou & Amygdalos, 2015). It was initiated by two patients with advanced ALS living in Brazil and the United States, both of whom died before the end of the study. ALS patients, 149 according to Vayena and colleagues (Streuli & Vayena, 2015; Vayena et al., 2016) or 160 according to Wicks (2018), were contacted via the platform and decided to take lithium to test the results of a small study that suggested positive effects in terms of slowing disease progression and alleviating symptoms. The ALS PatientsLikeMe study took place over eight months and was published in Nature Biotechnology. Its findings that lithium had no effect were later confirmed by standard clinical trials (Streuli & Vayena, 2015; Vayena et al., 2016; Wicks, 2018). However, the authors acknowledged the limitations of this type of self-experimentation, which raises a series of questions about the control of these trials, the lack of assistance when needed, and the risks faced by individuals. These initiatives also open up a debate regarding how “classical” research can be linked to what is called “patient-led research” in the literature.

---

3 PatientsLikeMe is an exchange platform for communities of patients, including patients suffering from chronic and rare diseases.
The involvement of patient organizations in the development and use of PROs is an important step in the development of tools that collect not only patient reports on questions developed by experts, but also information about patients’ experiences and their own assessments of what counts in new care strategies or treatments. The articulation of these PREMs or PC-PROs relative to classic PROs is currently an ongoing project, which will undoubtedly require work on a case-by-case basis.

Conclusion

PROs are data collected from patients about their own health status, particularly their health-related quality of life. It was in the context of the evaluation of new treatments tested in clinical trials that the FDA proposed the first formal definition of PROs in the 2000s. Since then, PROs have been used for a variety of purposes ranging from pharmacovigilance to alerts from patients on changes in their state of health, and from the collection of physiological data via individual digital devices and health platforms for clinical research to the development of medico-economic indicators to inform decisions on the management of health products by national authorities.

The subjective nature of patient reports has been the subject of intense reflections within communities of experts, and has led to methodological research aimed at standardizing the questions asked, developing graduated scales of response, and testing and validating the tools with different populations of concerned individuals. Another element of discussion relates to the conditions for the development of disease-specific PROs or particular dimensions of patients’ health status (e.g., pain, fatigue, etc.).

Recently, a new question has arisen regarding the ability of PROs to truly reflect patients’ experiences. Indeed, PROs ask questions related to subjects that experts consider important in their practices, but which are not necessarily relevant to patients. To address this issue, partnerships have been established with patient organizations to develop qualitative surveys that reveal the therapeutic benefits patients expect from a new drug in light of their experiences with the disease, or the problems that, from their point of view, impact their quality of life. These initiatives mainly concern chronic and rare diseases, and address issues such as the variety of symptoms or the expression of pain and fatigue.

To sum up, although PROs are now part of the health landscape, the literature shows that their development and articulation with other real-life data require a substantial amount of collaborative work and framing, and that concerns about their use remain, especially in France where few have developed expertise with PROs, unlike in the United States and the United Kingdom. In particular, the use of PROs in clinical practice is still very tentative, probably because it would require changing patient care practices and the organization of care.


ICHOM - International Consortium for Health Outcomes Measurement. (2018, circa). What is value-
based health care? Available on the ICHOM website: https://www.ichom.org/faqs/


Weldring, T., & Smith, S. M. S. (2013). Patient-reported outcomes (PROs) and patient-reported outcome measures (PROMs). Health Services Insights, 6, 61-68. https://doi.org/10.4137/HSI.S11093


