Received: 22 December 2022 | Revised: 27 April 2023 | Accepted: 3 May 2023

DOI: 10.1002/jimd.12622

REVIEW

JIMD SSIEM WILEY

# Measuring what matters: Why and how to include patient reported outcomes in clinical care and research on inborn errors of metabolism

**Martina Huemer**[1,2] | **Florin Bösch**[3,4]

[1]Division of Metabolism and Children's Research Center, University Children's Hospital Zurich, University of Zurich, Zurich, Switzerland

[2]Department of Paediatrics, LKH Bregenz, Bregenz, Austria

[3]Department of Psychosomatics and Psychiatry, and Children's Research Center, University Children's Hospital Zurich, University of Zurich, Zurich, Switzerland

[4]Division of Child and Adolescent Health Psychology, Department of Psychology, University of Zurich, Zurich, Switzerland

**Correspondence**
Martina Huemer, Division of Metabolism, University Children's Hospital Zurich, Steinwiessstrasse 75, 8032 Zurich, Switzerland and Landeskrankenhaus Bregenz, Dept. of Paediatrics, Carl-Pedenz-Str. 2, 6900 Bregenz, Austria.
Email: martina.huemer@kispi.uzh.ch; martina.huemer@lkhb.at

**Communicating Editor**: Peter Burgard

## Abstract

Patient reported outcomes (PROs) are generally defined as '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'. A broader definition of PRO also includes 'any information on the outcomes of health care obtained directly from patients without modification by clinicians or other health care professionals'. Following this approach, PROs encompass subjective perceptions of patients on how they function or feel not only in relation to a health condition but also to its treatment as well as concepts such as health-related quality of life (HrQoL), information on the functional status of a patient, signs and symptoms and symptom burden. PRO measurement instruments (PROMs) are mostly questionnaires and inform about what patients can do and how they feel. PROs and PROMs have not yet found unconditional acceptance and wide use in the field of inborn errors of metabolism. This review summarises the importance and usefulness of PROs in research, drug legislation and clinical care and informs about quality standards, development, and potential methodological shortfalls of PROMs. Inclusion of PROs measured with high-quality, well-selected PROMs into clinical care, drug legislation, and research helps to identify unmet needs, improve quality of care, and define outcomes that are meaningful to patients. The field of IEM should open to new methodological approaches such as the definition of core sets of variables including PROs to be systematically assessed in specific metabolic conditions and new collaborations with PRO experts, such as psychologists to facilitate the systematic collection of meaningful data.

**KEYWORDS**

health literacy, HrQol, MetabQol, outcome measures, patient education, patient-centred medicine

## 1 | PATIENT REPORTED OUTCOMES

Patient reported outcomes (PROs) are generally defined as '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'.[1] A broader definition of PRO also includes „any information on the outcomes of health care obtained directly from patients without modification by clinicians or other health care professionals'.[2] Following this approach, PROs encompass subjective perceptions of patients on how they function or feel not only in relation to a health condition but also to its treatment, as well as concepts such as health-related quality of life (HrQoL), information on the functional status of patients, disease signs and symptoms and symptom burden.[2] PRO measurement instruments (PROMs) are mostly questionnaires and inform about what patients can do and how they feel.[3]

## 2 | WHY ASSESS PROS IN PATIENTS WITH INBORN ERRORS OF METABOLISM?

Intriguingly, the meaning in terms of content of the proverb 'beauty is in the eye of the beholder' is known in many cultures around the world. It refers to the intuitive knowledge that each human being has a unique perception of the world. PROs help to bring the unique perspectives of patients to attention.

Although holistic care for chronically ill patients has been an integral part of research and clinical approach in the field for many decades PROs have not yet found as much entrance into clinical care and research related to inborn errors of metabolism (IEM) as in other chronic conditions, such as cancer or cystic fibrosis.[4,5]

While perceptions of health professionals regarding their patients are an extremely valuable source of information, they differ substantially from how patients perceive their situation and feel about it, and what is important and meaningful to them.[4] Assessments of medical conditions, severity of symptoms, and patients' psychological status by physicians is biased by different backgrounds of physicians and patients.[6] Quality and success of the communication between physicians and patients still depend highly on the degree of shared language, education, and ethnicity.[7] Thus, any thorough assessment of patients' situations or the effects of a treatment should include data provided by the patients themselves since doctor not always 'knows best', how patients feel and what they value.

PRO assessment helps identifying unmet needs as well as burdens of patients and caregivers related to a disease, and its treatment. Beyond burdens imposed by direct effects of a disease on physical, cognitive, or emotional performance, the rarity of a disease exerts specific additional strains on patients and families.[8] Patients with rare diseases such as IEM experience extended delays to diagnosis.[9] They may not easily find a specialist for their disease at all – and rather not in the vicinity of their homes.[9] When travelling and in need of medical help, they may be confronted with physicians unfamiliar with their disease. Treatment may not be available at all or access to it may be restricted.[8] Furthermore, if only small numbers of patients with a disease are known, patients experience that their prognosis and future disease course are uncertain and unforeseeable.[5,9]

In recent years, the further developments of newborn screening methods, treatment and care options have fostered a dramatic rise of the survival times of patients with IEM.[10] Therapies are often costly, and it is in the best interest of patients, physicians, and the health care system to understand whether the expectations regarding their effect are met for all sides.[2] Treatment effects may be difficult to estimate with a solely 'physical' approach in small cohorts of patients.[11–13] Traditional surrogate markers such as phenylalanine concentrations in PKU, or liver size in Gaucher disease give proof that a treatment does what it promises to do in terms of reducing harmful substances or clearing of cells from storage material. These data are incredibly important and a prerequisite of any research-based treatment, but they cannot inform about how patients' feel their lives to be changed for the better, the worse or not at all by a treatment. To gain this information, PROs need to be included into research and care settings.[14]

Inclusion of PROs is especially important in conditions with a variable presentation and natural course, in which treatment finds individuals in very different states of the condition, or if treatment can only slow down clinical deterioration in a progressive disease. Both scenarios make the estimation of treatment benefits extremely complex. In such circumstances, the patients' perspective on the change a treatment makes for them in everyday life an important complementary source of information to answer the question whether a treatment can be considered appropriately chosen, and successful so that it should be further supported and reimbursed.[2,4]

PROs are not only valuable in research environments or regulatory circumstances such as drug development and approval, but also in daily clinical practice. Self-reports help to identify less well-known or unexpected symptoms of a disease, emotional burdens of patients, and side effects of treatments. The use of self-reporting

PROs in clinical care of cancer patients has a positive effect on the number of emergency admissions as well as on quality-adjusted survival.[15] PRO data inform and support clinical decision-making, for example, by allowing to weigh subjective benefits against side effects of a treatment,[16] and by facilitating communication between clinicians and patients.[5] Furthermore, PROs may support the growing interest in precision medicine, an approach focused on providing the 'right' diagnostic tools and treatments for each individual patient. PROs and psychosocial factors in general may in this interest complement information from genetic, and biomarker parameters.[17]

Inclusion of PROs not only improves patient satisfaction, disease- and symptom management skills, and HrQoL[15,18,19] but makes targeted interventions possible.[5]

For example, can health literacy be improved by targeted interventions. Health literacy is 'the degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions.[20,21] Patients with IEM and their caregivers report that they often find it difficult to understand the pathophysiology of the disease and the rationale of its treatment. It is hard to understand the complex genetic background as well as intracellular pathophysiological processes such as an enzyme that is not properly operating. Furthermore, signs of a disease may evolve in subtle steps that cannot easily be observed in daily life.[22] Parents and patients find it hard to explain the disease to others when searching social support.[22] Our group has developed, and tested patient education materials that were presented in standardised single or group training sessions to 74 controls and 37 patients with IEM (phenylketonuria, galactosemia, urea cycle defects, lysosomal storage disorders). The gain of disease-specific health literacy post intervention was highly significant in both groups.[23] Since limited health literacy is associated with less successful outcomes in chronic disease, more medication errors, and more hospitalisations, improvement of health literacy is a meaningful target to improve physical outcomes and other PRO.[24–26]

Another example for targeting and evaluating interventions is the observation that HrQol in children and adolescents with chronic diseases is determined by their coping patterns (effective vs. non-effective coping). Improvements of HrQol can be achieved by interventions targeting coping behaviour.[27]

From the perspective of the health system, PROs are not only useful to identify and target risk factors for limited treatment success (such as low health-literacy or ineffective coping) and general outcome but to bridge boundaries between disciplines and services, guide interventions and monitoring of their effects, and to identify needs for adaption of health strategies and policies, legislation, and resource allocation.[28] In the circumstances of drug or medical product approval, trials with medical and biochemical parameters as primary outcomes profit from the inclusion of PROs that provide additional, decision-relevant evidence.[2]

## 3 | PATIENT REPORTED OUTCOME MEASURES (PROMS)

PROMs are mostly paper or computer-based self-completed questionnaires. Self-reporting is the optimal PRO assessment method, but observer-reported outcomes (ObsROs or proxy-reports) measured by observer-reported outcome measurement (ObsROMs) tools are widely used as substitute for self-reports in young children, or in patients unable to answer such questionnaires. ObsROs are reported by a person that is close to the patient in daily life. Mostly, observers are parents or caregivers.[29] The degree of concordance of PRO and ObsRo data depends of course predominantly on the external perceivability of what is measured. A sign of a disease such as fever will probably be reported more concordantly than discomfort associated with fever.[30,31] The degree of concordance is also influenced by patient age,[32] illness severity,[33] and observer characteristics such as how distressed the observer feels.[34]

Beyond being a substitute for self-reported data, ObsROs are a valuable complementary source of information in specific settings.[35] Parents may provide valuable insights into their child's behaviour or disease management skills.[36] It has thus been recommended to include perspectives of patients and observers, and to explore the relationship between the two to understand more about patients and their daily experiences.[37] The exploration of relations between ObsROs and self-reported PROs in children and adolescents with acute intoxication type IEM and PKU revealed that the more severe the parents perceive their child's condition to be (a parents' self-report on their perception), the higher was their psychological burden and the lower were patient self- and parent reported (ObsRo) HrQoL scores. This example illustrates that it is worthwhile to explore relations between parents' self- assessments, PROs and ObsROs because targeting the perception of the severity of my child's condition (that is not necessarily correlated with severity assessments by metabolic specialists) by an intervention can be the key to improving self- and parent reported HrQol of the patients.[38]

Whether a PROM chosen for clinical care or research is generic or disease specific determines the information it can provide. A generic PROM captures widely applicable concepts such as limitations of physical activities,

energy and emotion and their impact on social activities and is informative on the burden patients experience in comparison to the general population or to patients with other health conditions. Frequently used generic instruments have often been validated in multiple languages and populations.[2,4]

Disease-specific instruments target concepts directly associated with a condition,[35] such as the impact of the disease and dietary treatment on HrQol[39] or the burden of illness[40] in patients with PKU, or a in group of conditions sharing essential characteristics such as intoxication type IEM,[41] or primary mitochondrial diseases.[42] Variations of the severity of disease-specific symptoms in lysosomal storage disorders may be assessed by severity scores that are complemented by observers.[43] By addressing disease-specific concepts, PRO assessment becomes more informative for clinicians and researchers. Although it has recently been recommended to develop specific PROMs for patients with IEM,[42,44] such tools are unfortunately still scarce. Furthermore, existing disease-specific PROMs may also be largely unknown among researchers and clinicians, or unavailable to them due to complex or costly licencing procedures. In a scoping review on the use of PROMs for caregivers and patients with IEM including 131 studies, that applied 32 HrQol instruments; only two of these instruments were disease specific (PKU-QoL[39] and QoL Scale for Metabolic Diseases[45]).[46]

The construction of a disease-specific PROM is a complex, time- and work-intense process. The first step is to identify representatives for the targeted populations, for example, patients with IT-IEM and their families, experts such as nurses, dieticians, psychologists, and physicians involved in the care for this patient group from all relevant sociocultural backgrounds.[47] From this group, contents are collected, mostly by single interviews or focus groups. Statements from the group are transcribed and undergo qualitative analysis during which they are categorised according to their contents and ordered according to how often and with how much emphasis they have been presented. Based on the list of contents, a preliminary set of questionnaire items is developed and presented to the target population for cognitive debriefing, a process in which it is evaluated whether participants understand the items, find any contents missing or represented incomprehensively. This feedback is used to adapt the items and assemble them for a first questionnaire that consecutively undergoes psychometric evaluation.[14]

In diseases with a very heterogeneous presentation such as mucopolysaccharidosis type 1 or type 2 with a clinical spectrum from carpal tunnel syndrome and frequent infections of the upper respiratory tract to severe, progressive cognitive impairment, osteopathy, coarse facial features and enlarged organs it may be impossible to find a smallest common of questions that make sense for all patients. For such heterogeneous populations it is worth discussing whether the nomothetic approach that focuses on differences between individuals with a specific disease is more helpful compared to assessing the variation of individual patients' characteristics over time or the idiosyncratic approach. Combinations of both approaches using mathematical algorithms have shown promising results, for example, in the psychiatric field.[48] Goal attainment scales may also be introduced to assess how patients are doing in achieving their personal health-related goals and wishes.[4]

PROMs may be developed according to the classical test theory. Here, series of items are grouped according to content or dimensions and define subscores or summative scores. The 'Patient Reported Outcomes Measurement Information System' or PROMIS initiative has created PROMs based on the item response theory. Item pools targeting a specific content, or 'latent trait' are ordered by a gradient such as severity. For example, can the item 'are you able to move independently between rooms in the house' be considered a 'severe' item to address exercise intolerance while 'can you walk uphill' constitutes a 'less severe' item. IRT-based instruments can be administered as fixed forms. In addition to this, PROMIS has developed Computer Adaptive Tests (CATs) in which the first item presented to the respondent is usually from about the middle of the severity range. Each next item presented to the respondent is selected on basis of the accumulating answers to earlier items. During the process, the profile of an individual is more and more 'sculpted' and refined.

PROMIS has been initiated by the National Institute of Health in the United States, and in recent years, PROMIS measures have been translated into many other languages for use across the world. The PROMIS approach has high potential to improve research and patient care, also in the field of IEM.[49]

## 4 | QUALITY CRITERIA FOR PROMS

'Selecting unsuitable or insufficient quality outcome measurement instruments may introduce bias in the conclusions of studies. This may lead to a waste of resources and is unethical because participating patients contribute little or nothing to the body of evidence but still suffer from the burdens and risks of the study'.[50] This statement from the COSMIN (COnsensus-based Standards for the selection of health Measurement INstruments) initiative makes it more than clear that to collect high-quality data in an ethically responsibly manner, it is mandatory

to use high-quality PROMs. The development, modification, and selection of PROMs require as much care as is invested in the choice of physical or biochemical assessment methods.

PROMs must be standardised, reliable, valid, and should generate quantifiable data. Completing them should not be extensively burdensome or time-consuming.[4]

Standardised means that an instrument is completed and evaluated in a defined way. Reliability informs about the degree to which an instrument is free from measurement errors and depicts changes that are real and not attributable to its poor construction and performance.[2] Reliability specifies the consistency of scores over time intervals, during which changes are not expected (reproducibility). If multi-item scales are used, the internal consistency coefficient quantifies the extent to which items of shared content confirm each other.[2,16] Inter-rater reliability of an instrument is of relevance when scores are provided by several observers or clinicians to rate, for example, the clinical symptoms of patients.[16]

Validity informs about how accurate an instrument measures what it claims to measure.

Content validity is the degree to which the content of a PROM is important, meaningful, and comprehensive for the investigated population (e.g., patients, parents), the situation, the research question, or specific clinical circumstances.[2] If a well-established instrument exists that is supposed to measure a certain construct (such as HrQoL), there should be a sufficient correlation with other PROMs claiming to measure the same construct.

If, for example a group of patients with acute intoxication type IEM (organic acidurias, maple syrup urine disease, urea cycle disorders) are known or highly expected to score differently on the burden of emergency hospital admissions compared to patients with PKU, who have no acute metabolic crises, the PROM should depict these differences as well (construct and criterion validity).[2]

A most important characteristic of a PROM in research and clinical care is its ability to detect changes (responsiveness) meaning that it depicts changes a new treatment or other intervention is expected to induce.[2]

Generic instruments are often very well tested for reliability and validity but are usually less responsive to change following an intervention or modification of treatment for a given disease than specific instruments. Such changes can, although subtle, be meaningful and important to patients and should be captured by a responsive PROM to properly evaluate what a treatment changes in patients' lives. While a main general limitation of disease-specific PROMs for IEM is that they are less well validated due to the limited numbers of patients, or heterogeneous disease presentations, disease-specific instruments are considered more responsive.[4]

Weighing the advantages and shortcomings of generic and specific PROMs makes it advisable to combine generic and disease specific PROMs to capture patients' status quo more broadly and to be able to observe meaningful changes.[2] For example, can HrQol in children with urea cycle disorders be studied by combining the age-appropriate versions of the generic PedsQl self-reporting scales with proxy-reporting generic PedsQl scales for parents (an ObsROM) and the respective self- and proxy-reporting versions of the disease-specific PROM MetabQoL that has been specifically developed for patients with intoxication type IEM and focuses on aspects not covered by a generic questionnaire.[41] Using this approach, PROMs allow for self-reporting, ObsROMs inform about very young or cognitively impaired patients. The generic PedsQl assesses the general concept of HrQoL and allows for comparison, for example, with patients following a diet for PKU or diabetes, or the general population. The specific MetabQol contributes information of disease-related burdens such as fear of metabolic crises and emergency admissions or the impact of treatment with amino acid supplements and limitation of natural protein intake.[41]

Talking about HrQol, the probably most often used outcome parameter in chronic disease in the intent to assess patients' overall perception of the impact of an illness and its treatment,[3] it must be considered that HrQol is not a simple parameter but a complex, multidimensional psychological construct, usually encompassing the three major dimensions physical, mental (emotional and cognitive), and social well-being.[2,14] The dimensions are measured by subscales of questions that generate summary scores for each dimension. These subscores define a profile of HrQol (e.g., good physical but impaired social well-being) and if added to a total score, quantify general HrQol.[14] The relevance of the subdomains of HrQol may vary across diseases. A newly developed disease specific HrQol instrument will have content validity but may address subdomains deviating from the above-mentioned general dimensions of HrQol.[37,51] It is mandatory to clarify for each adapted or newly developed HrQol instrument which dimensions it addresses to make any concept alterations transparent in the interest of avoiding subliminal shifts towards different definitions of HrQol[14] and ending up with comparing apples and oranges.

## 5 | HOW TO SELECT PROS AND PROMS IN CLINICAL CARE AND RESEARCH?

PROs encompass reports of the status of a patient's health condition measured directly from the patient. Some are clear, situational assessments such as the assessment of

intensity and time of procedural-associated pain or nausea related to medication intake. Others, such as HrQoL or disease knowledge are more complex constructs.[28] The selection of PROs that are of interest in each disease, population, and situation and meaningful to clinical or research questions, or claims of new treatments during regulatory processes requires most careful attention and specific expertise.

First, the ideas of what concepts may be of interest need to be broken down to outcomes that are formulated unequivocally and tested for their congruency with the target. What claims shall be made for a new drug and be supported by PROs? Which PROs are explored regarding their responsiveness to a psychological patient training? How are PROs of interest interconnected? For example, general health behaviour or attitudes towards health care are PROs with an impact on other PROs such as self-reported adherence to treatment or HrQoL.[3] The more meticulously the outcomes are defined, the easier it becomes to select the appropriate PROMs.

In children and adolescents with Fabry disease, outcomes of interest could for example be the intensity of procedure-associated pain in situations of intravenous enzyme replacement therapy as well as intensity, duration, and number of neuropathic pain attacks.[52] While both PROs are about pain, their qualities and situational aspects are substantially different and require different PROMs to measure them.[51] In PKU patients, outcomes of interest could be the impact of dietary restrictions and supplement intake on HrQol that can be evaluated using the disease specific PKU-QoL.[39] The impact of daily injections of the enzyme replacement pegvaliase that may cause severe allergic reactions and thus requires precautions would require a specific PROMs approach as it is not covered by the PKUQoL.[39]

Some excellent databases and collections of PROMs that have been assembled in recent years such as PROMIS,[49] the National Insitute of Health (NIH) toolbox,[53] the COSMIN[50] materials and a collection of PROMs specifically assembled for IEM patients[54] facilitate the choice of PROMs relevant for a specific outcome in a defined population.

Because it is probably illusionary and economically impossible to develop new, specific PROMs for each of the more than 1000 IEM known today, it is also reasonable to adapt existing PROMs to a new disease and population. Parts of the PKUQOoL could probably be adapted for classical homocystinuria because the diseases share characteristics of treatment (protein-restricted diet with amino acid supplements prevents severe symptoms) although being clinically very different. For this approach the quality criteria as outlined above of course apply, too. When adapting an existing tool, an abbreviated and facilitated development process for a new instrument is possible but content validity requires specific attention. Obtaining direct input from patients with the 'new' disease on contents covered and not yet covered, as well as wording of the existing PROMs by qualitative studies (interviews, focus groups) are still required.[55]

## 6 | INTERPRETATING PROMS DATA

Numeric results retrieved by PROMs can be statistically analysed and related to other data on patients such as, for example, the biochemical response to a drug, or reduction of seizure frequency.[2,56] An additional approach is the use of thresholds such as the minimally important changes (MIC) of PROMs scores. MIC must be differentiated from minimal detectable differences, which are scores that can be assessed by statistical methods with a reasonable degree of certainty.[56]

MIC build on the individual patient's attribution of the minimal 'meaningful' – and thus important – change over time. In other words, the MIC is the minimal change that makes an individual patient feel changed for the better or the worse over time. MIC are established using anchor questions. A PROM score is related to patients' assessment on how things have changed for them (e.g., on a 5-point Likert scale from 'much worse' to 'much better'). The mean of individual MICs or the number of patients that can be considered responders based on them reaching their individual MIC following an intervention can be useful MIC-related interpretation tools.[56]

In adults with Pompe disease, a lysosomal storage disease affecting the muscle and progressive despite a very costly intravenous enzyme replacement therapy, the 6-min walk test and forced vital capacity are the most frequently used outcomes to test treatment efficacy. It has been discussed controversially whether a significant change from baseline is sufficient to indicate treatment success or whether it must also be investigated to what extent the change is meaningful for patients. Patient's comments on whether they felt their abilities 'unchanged', or found the changes 'marginal', 'more than minimal' or 'significant' were used to anchor the quantitative results of forced vital capacity and 6-min walk test. The MIC were used to retrospectively explore, whether enzyme replacement therapy studies in late-onset Pompe disease resulted in changes in both tests exceeding the MIC. The study has the major limitation that the MIC data were transferred from a different population (patients suffering from other chronic respiratory diseases) and not validated for the Pompe group but the general approach to relate numeric results to patients' perceptions should be further pursued in IEM.[57]

## 7 | INTEGRATION OF PROS/PROMS INTO CLINICAL ROUTINE AND RESEARCH ON IEM

Ideally, PROs should be selected by a group of specialists and patients/caregivers as part of a core-set of parameters recommended to be used as outcomes in research, drug legislation studies, and clinical care for a given disease. This approach would make the development of more disease-specific PROMs worthwhile and allow for more uniform study designs that have a higher probability of collecting more homogeneous data from larger populations.[58]

In juvenile idiopathic arthritis, a group of diseases comparable to IEM regarding their heterogeneity and rarity, core sets of parameters to be considered for the evaluation of children with JIA in research and clinical care have been in place for many years.[58,59] Patients, caregivers, patient representatives as well as clinicians and researchers in the field contributed to the selection of the core set domains using consensus methods.[60] The domains encompass physicians' clinical assessments, PROs and ObsRos, as well as laboratory test results.[58] A similar approach could be extremely helpful for research and clinical care for patients with IEM.

For conditions that are extremely heterogeneous and/or progressive, the treat-to-target approach that is closely related to the concepts of MIC responders and goal attainment scales may be especially useful.[4] The treat-to-target approach has been successfully implemented in diabetes,[61] Crohn's disease,[62] and paediatric rheumatology[59] care and research. It can be applied to PROMs as well as to biochemical or physical parameters and is based on identification and definition of a core set of appropriate treatment targets under consideration of the available evidence and defined in a shared decision process involving experts, patients, and caregivers. Interventions are tested regarding their performance to reach these pre-defined targets.[62]

## 8 | CLINICAL AND RESEARCH PERSPECTIVES FOR PROS/PROMS IN IEM

Even if metabolic experts consider PROs and PROMs important, they often cannot assess them due to limited time resources, unavailability of PROMs, and absence of PROs/PROMs experts in the metabolic teams (unpublished data from a survey within the European Network and Registry for Homocystinurias and Methylation Defects {EHOD} project). It must be considered, however, that if PROs/PROMs are not included in care and research settings, their present shortcomings such as not being validated for a specific population of interest, being unavailable in a certain language, or being cumbersome and expensive to retrieve, will sustain. As with biochemical methods, the only way of improving PRO-related methods is their use and continuous improvement to allow – with time – for more patient data that reflect the

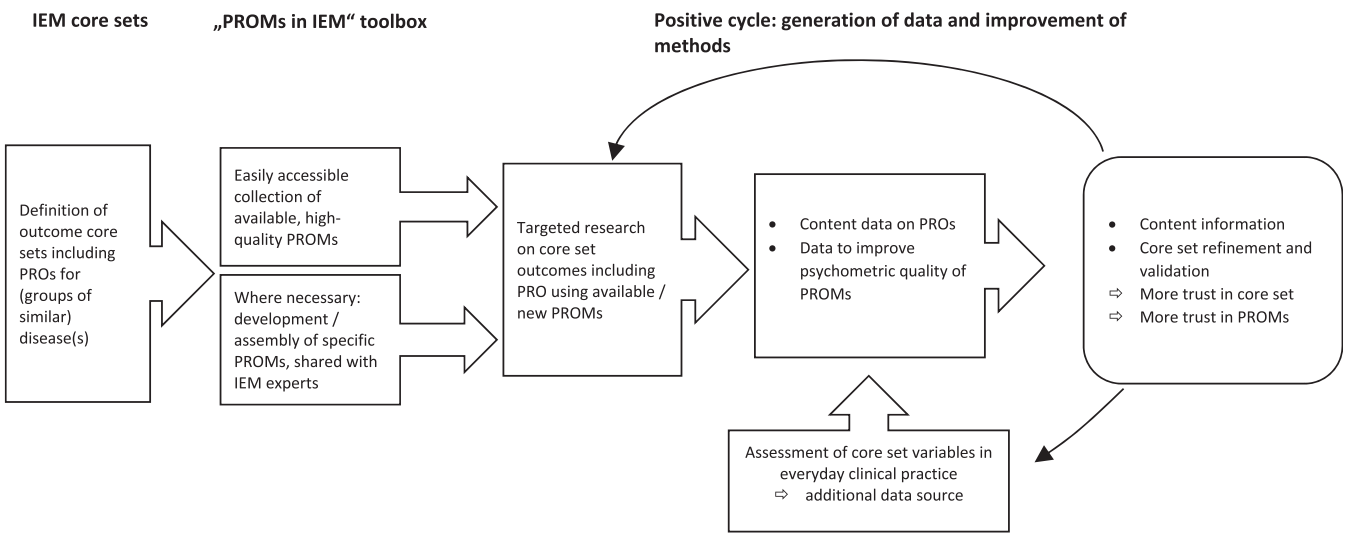


**FIGURE 1** Suggestions for a strategy for the metabolic community to include PROs into outcome core sets for inborn errors of metabolism: a defined core set of important outcome variables for IEM in research and clinical settings including PROs, and use of easily available, high-quality PROMs facilitate acquisition of important clinical data and data that is required to refine and improve PROMs. IEM, inborn errors of metabolism; PROs, patient reported outcomes; PROMs, PRO measurements.

unique perspective of those that all our care is about – our patients and their families.

Standardised approaches towards PROs/PROMs for research, clinical care and drug legislation processes could be advantageous and foster the use of PROs in the IEM field. PROs should be integral part of outcome core sets for IEM, and the metabolic community should have access to suitable PROMs. Following such a strategy would not only generate important clinical data but also data to further improve the psychometric qualities of PROMs (Figure 1).

## 9 | CONCLUSIONS AND PERSPECTIVES

Inclusion of PROs assessed using high-quality, well-selected PROMs into clinical care, legislative, and research settings helps to identify unmet needs, improve quality of care, and define outcomes that are meaningful to patients. The field of IEM should open itself to new methodological approaches such as the definition of core sets of variables to be systematically assessed in specific metabolic conditions and new collaborations with PRO experts, such as psychologists to facilitate the systematic collection of meaningful data.

### ACKNOWLEDGEMENTS

We thank Marion Herle, Markus Landolt, and Matthias R. Baumgartner for fruitful discussions and their ongoing support. Open access funding provided by Universitat Zurich.

### CONFLICT OF INTEREST STATEMENT

Martina Huemer and Florin Bösch declare no conflicts of interest regarding the contents of this manuscript. M.H. has in the past 5 years received unrestricted research grants from Nutricia Metabolics, and Sanofi Genzyme as well as honoraria for lectures or scientific advice from Nutricia Metabolics, Recordati Rare Disease Foundation, and Aeglea. Florin Bösch has been supported by Vontobel-Stiftung, Zürich, Switzerland.

### ORCID

*Martina Huemer* https://orcid.org/0000-0002-0590-678X

### REFERENCES

1. US Department of Health and Human Services Food and drug administration: patient-reported outcome measures: use in medical product development to support labeling claims. Guidance for Industry; 2009. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf. Accessed November 20, 2022
2. Cella D, Hahn EA, Jensen SE, et al. *Patient-Reported Outcomes in Performance Measurement*. RTI Press; 2015.
3. Patrick DL, Burke LB, Powers JH, et al. Patient-reported outcomes to support medical product labeling claims: FDA perspective. *Value Health*. 2007;(10 Suppl 2):S125-S137.
4. Rüther A, Elstein D, Wong-Rieger D, Guyatt G. Aspects of patient reported outcomes in rare diseases: a discussion paper. *Int J Technol Assess Health Care*. 2016;32(3):126-130. doi:10.1017/S0266462316000271
5. Slade A, Isa F, Kyte D, et al. Patient reported outcome measures in rare diseases: a narrative review. *Orphanet J Rare Dis*. 2018;13(1):61. doi:10.1186/s13023-018-0810-x
6. Zola IK. Problems of communication, diagnosis, and patient care: the interplay of patient, physician and clinic organization. *J Med Educ*. 1963;38:829-838.
7. Aelbrecht K, Hanssens L, Detollenaere J, Willems S, Deveugele M, Pype P. Determinants of physician-patient communication: the role of language, education and ethnicity. *Patient Educ Couns*. 2019;102(4):776-781. doi:10.1016/j.pec.2018.11.006
8. Benjamin K, Vernon MK, Patrick DL, Perfetto E, Nestler-Parr S, Burke L. Patient-reported outcome and observer-reported outcome assessment in rare disease clinical trials: an ISPOR COA emerging good practices task force report. *Value Health*. 2017;20(7):838-855. doi:10.1016/j.jval.2017.05.015
9. Lagler FB, Moder A, Rohrbach M, et al. Extent, impact, and predictors of diagnostic delay in Pompe disease: a combined survey approach to unveil the diagnostic odyssey. *JIMD Rep*. 2019;49(1):89-95. doi:10.1002/jmd2.12062
10. Almannai M, Marom R, Sutton VR. Newborn screening: a review of history, recent advancements, and future perspectives in the era of next generation sequencing. *Curr Opin Pediatr*. 2016;28(6):694-699. doi:10.1097/MOP.0000000000000414
11. Spillmann RC, McConkie-Rosell A, Pena L, et al. A window into living with an undiagnosed disease: illness narratives from the undiagnosed diseases network. *Orphanet J Rare Dis*. 2017;12:71.
12. EURORDIS. The voice of 12,000 patients: experiences and expectations of rare disease patients on diagnosis and Care in Europe: EURORDIS rare diseases. *Europe*. 2009; https://www.eurordis.org/wp-content/uploads/2009/12/EURORDISCARE_FULLBOOKr.pdf. Accessed November 20, 2022
13. Angelis A, Tordrup D, Kanavos P. Socio-economic burden of rare diseases: a systematic review of cost of illness evidence. *Health Policy*. 2015;119:964-979.
14. Bullinger M, Quitmann J. Quality of life as patient-reported outcomes: principles of assessment. *Dialogues Clin Neurosci*. 2014 Jun;16(2):137-145. doi:10.31887/DCNS.2014.16.2/mbullinger
15. Basch E, Deal AM, Kris MG, et al. Symptom monitoring with patient reported outcomes during routine cancer treatment: a randomized controlled trial. *J Clin Oncol*. 2016;34:557-565.
16. Kim HS, Courtney DM, McCarthy DM, Cella D. Patient-reported outcome measures in emergency care research: a primer for researchers, peer reviewers, and readers. *Acad Emerg Med*. 2020;27(5):403-418. doi:10.1111/acem.13918
17. Ramaswami R, Bayer R, Galea S. Precision medicine from a public health perspective. *Annu Rev Public Health*. 2018;39:153-168. doi:10.1146/annurev-publhealth-040617-014158
18. Hjollund INH, Larsen PL, Biering K, Johnsen PS, Riiskjær E, Schougaard ML. Use of patient-reported outcome (PRO)

measures at group and patient levels: experiences from the generic integrated PRO system. *WestChronic Interact J Med Res*. 2014;3:e5.
19. Jensen RE, Rothrock NE, DeWitt EM, et al. The role of technical advances in the adoption and integration of patient-reported outcomes in clinical care. *Med Care*. 2015;53:153-159.
20. US Department of Health and Human Services. *Healthy People 2010: Understanding and Improving Health*. 2nd ed. US Government Printing Office; 2000.
21. Okuhara T, Ishikawa H, Okada M, Kato M, Kiuchi T. Designing persuasive health materials using processing fluency: a literature review. *BMC Res Notes*. 2017;10(1):198. doi:10.1186/s13104-017-2524-x
22. Baker DW, Gazmararian JA, Williams MV, et al. Functional health literacy and the risk of hospital admission among Medicare managed care enrolees. *Am J Public Health*. 2002;92(8):1278-1283.
23. Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K. Low health literacy and health outcomes: an updated systematic review. *Ann Intern Med*. 2011;155:97-107.
24. DeWalt DA, Berkman ND, Sheridan S, Lohr KN, Pignone MP. Literacy and health outcomes: a systematic review of the literature. *J Gen Intern Med*. 2004;19(12):1228-1239.
25. Zeltner NA, Landolt MA, Baumgartner MR, et al. Living with intoxication-type inborn errors of metabolism: a qualitative analysis of interviews with Paediatric patients and their parents. *JIMD Rep*. 2017;31:1-9. doi:10.1007/8904_2016_545
26. Zeltner NA, Welsink-Karssies MM, Landolt MA, et al. Reducing complexity: explaining inborn errors of metabolism and their treatment to children and adolescents. *Orphanet J Rare Dis*. 2019;14(1):248. doi:10.1186/s13023-019-1236-9
27. Oppenheimer S, Krispin O, Levy S, Ozeri M, Apter A. The impact of coping patterns and chronic health conditions on health-related quality of life among children and adolescents. *Eur J Pediatr*. 2018;177(6):935-943. doi:10.1007/s00431-018-3146-6
28. Weldring T, Smith SM. Patient-reported outcomes (PROs) and patient-reported outcome measures (PROMs). *Health Serv Insights*. 2013;6:61-68. doi:10.4137/HSI.S11093
29. Basch E, Bennett AV. Patient-reported outcomes in clinical trials of rare diseases. *J Gen Intern Med*. 2014;29(3):801-804. doi:10.1007/s11606-014-2892-z
30. Andresen EM, Vahle VJ, Lollar D. Proxy reliability: health-related quality of life (HRQoL) measures for people with disability. *Qual Life Res*. 2001;10(7):609-619.
31. Matziou V, Perdikaris P, Feloni D, Moshovi M, Tsoumakas K, Merkouris A. Cancer in childhood: children's and parents' aspects for quality of life. *Eur J Oncol Nurs*. 2008;12(3):209-216. doi:10.1016/j.ejon.2007.10.005
32. Agnihotri K, Awasthi S, Singh U, Chandra H, Thakur S. A study of concordance between adolescent self-report and parent-proxy report of health-related quality of life in school-going adolescents. *J Psychosom Res*. 2010;69(6):525-532. doi:10.1016/j.jpsychores.2010.03.011
33. Morrow AM, Hayen A, Quine S, Scheinberg A, Craig JC. A comparison of doctors', parents' and children's reports of health states and health-related quality of life in children with chronic conditions. *Child Care Health Dev*. 2012;38(2):186-195. doi:10.1111/J.1365-2214.2011.01240.X
34. Galloway H, Newman E, Miller N, Yuill C. Does parent stress predict the quality of life of children with a diagnosis of ADHD? A comparison of parent and child perspectives. *J Atten Disord*. 2019;23(5):435-450. doi:10.1177/1087054716647479
35. Cella D, Hahn E, Jensen S, et al. Patient-reported outcomes in performance measurement. *RTI Press*. 2015. doi:10.3768/RTIPRESS.2015.BK.0014.1509
36. Matza LS, Patrick DL, Riley AW, et al. Pediatric patient-reported outcome instruments for research to support medical product labeling: report of the ISPOR PRO good research practices for the assessment of children and adolescents task force. *Value Health*. 2013;16(4):461-479. doi:10.1016/j.jval.2013.04.004
37. U.S. Department of Health and Human Services Food and Drug Administration. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. *Health Qual Life Outcomes*. 2006;4(1):1-20. doi:10.1186/1477-7525-4-79
38. Bösch F, Landolt MA, Baumgartner MR, et al. Caregiver burden, and parents' perception of disease severity determine health-related quality of life in paediatric patients with intoxication-type inborn errors of metabolism. *Mol Genet Metab Rep*. 2022;31:100876. doi:10.1016/j.ymgmr.2022.100876
39. Regnault A, Burlina A, Cunningham A, et al. Development and psychometric validation of measures to assess the impact of phenylketonuria and its dietary treatment on patients' and parents' quality of life: the phenylketonuria—quality of life (PKU-QOL) questionnaires. *Orphanet J Rare Dis*. 2015;10:59. doi:10.1186/s13023-015-0261-6
40. Quinn J, Georgiadis A, Lewis HB, Jurecki E. Measuring burden of illness in phenylketonuria (PKU): development of the PKU symptom severity and impacts scale as a robust patient-reported outcome. *Adv Ther*. 2022;39(2):971-991. doi:10.1007/s12325-021-01986-2
41. Zeltner NA, Baumgartner MR, Bondarenko A, et al. Development and psychometric evaluation of the MetabQoL 1.0: a quality of life questionnaire for Paediatric patients with intoxication-type inborn errors of metabolism. *JIMD Rep*. 2017;37:27-35. doi:10.1007/8904_2017_11
42. Mancuso M, McFarland R, Klopstock T, Hirano M, Consortium on Trial Readiness in Mitochondrial Myopathies. International workshop: outcome measures and clinical trial readiness in primary mitochondrial myopathies in children and adults. Consensus Recommendations. 16–18 November 2016, Rome, Italy. *Neuromuscul Disord*. 2017;27(12):1126-1137. doi:10.1016/j.nmd.2017.08.006
43. Evans W, Patterson M, Platt F, Guldberg C, Mathieson T, Pacey J. Core working Group for the Delphi Study. International consensus on clinical severity scale use in evaluating Niemann-pick disease type C in paediatric and adult patients: results from a Delphi study. *Orphanet J Rare Dis*. 2021;16(1):482. doi:10.1186/s13023-021-02115-6
44. Salek MS, Ionova T, Johns JR, Oliva EN. Appraisal of patient-reported outcome measures in analogous diseases and recommendations for use in phase II and III clinical trials of pyruvate kinase deficiency. *Qual Life Res*. 2019;28(2):399-410. doi:10.1007/s11136-018-2025-y
45. Eminoglu TF, Soysal SA, Tumer L, Okur I, Hasanoglu A. Quality of life in children treated with restrictive diet for inherited metabolic disease. *Pediatr Int*. 2013;55:428-433.
46. Pascoal C, Brasil S, Francisco R, et al. Patient and observer reported outcome measures to evaluate health-related quality of life in inherited metabolic diseases: a scoping review.

*Orphanet J Rare Dis*. 2018;13(1):215. doi:10.1186/s13023-018-0953-9
47. Boynton PM, Wood GW, Greenhalgh T. Hands-on guide to questionnaire research. Reaching beyond the white middle classes. *BMJ*. 2004;328:1433-1436.
48. Beltz AM, Wright AG, Sprague BN, Molenaar PC. Bridging the nomothetic and idiographic approaches to the analysis of clinical data. *Assessment*. 2016;23(4):447-458.
49. https://commonfund.nih.gov/promis/index.
50. https://database.cosmin.nl/.
51. www.cosmin.nl/finding-right-tool/deciding-outcome-measure/.
52. Burand AJ Jr, Stucky CL. Fabry disease pain: patient and preclinical parallels. *Pain*. 2021;162(5):1305-1321. doi:10.1097/j.pain.0000000000002152
53. https://www.healthmeasures.net/explore-measurement-systems/nih-toolbox.
54. Bösch F, Zeltner NA, Baumgartner MR, Huemer M, Landolt MA. Key patient-reported outcomes in children and adolescents with intoxication-type inborn errors of metabolism: an international Delphi-based consensus. *Orphanet J Rare Dis*. 2022;17(1):26. doi:10.1186/s13023-022-02183-2
55. Rothman M, Burke L, Erickson P, Leidy NK, Patrick DL, Petrie CD. Use of existing patient-reported outcome (PRO) instruments and their modification: the ISPOR good research practices for evaluating and documenting content validity for the use of existing instruments and their modification PRO task force report. *Value Health*. 2009;12(8):1075-1083. doi:10.1111/j.1524-4733.2009.00603.x
56. Terwee CB, Peipert JD, Chapman R, et al. Minimal important change (MIC): a conceptual clarification and systematic review of MIC estimates of PROMIS measures. *Qual Life Res*. 2021 Oct;30(10):2729-2754. doi:10.1007/s11136-021-02925-y
57. Lachmann R, Schoser B. The clinical relevance of outcomes used in late-onset Pompe disease: can we do better? *Orphanet J Rare Dis*. 2013;8:160.
58. van Dijkhuizen EHP, Ridella F, Naddei R, et al. Pediatric rheumatology international trials organization (PRINTO). Validity and reliability of four parent/patient-reported outcome measures for juvenile idiopathic arthritis remote monitoring. *Arthritis Care Res (Hoboken)*. 2022;75:391-400. doi:10.1002/acr.24855
59. Ravelli A, Consolaro A, Horneff G, et al. Treating juvenile idiopathic arthritis to target: recommendations of an international task force. *Ann Rheum Dis*. 2018;77(6):819-828. doi:10.1136/annrheumdis-2018-213030
60. Morgan EM, Munro JE, Horonjeff J, et al. Establishing an updated core domain set for studies in juvenile idiopathic arthritis: a report from the OMERACT 2018 JIA workshop. *J Rheumatol*. 2019;46:1006-1013.
61. Garber AJ. Treat-to-target trials: uses, interpretation and review of concepts. *Diabetes Obes Metab*. 2013;16:193-205.
62. Garcia NM, Cohen NA, Rubin DT. Treat-to-target and sequencing therapies in Crohn's disease. United European. *Gastroenterol J*. 2022;10:1121-1128. doi:10.1002/ueg2.12336

**How to cite this article:** Huemer M, Bösch F. Measuring what matters: Why and how to include patient reported outcomes in clinical care and research on inborn errors of metabolism. *J Inherit Metab Dis*. 2023;1-10. doi:10.1002/jimd.12622