Using Patient Experience Data to Evaluate Medical Interventions

Generating, understanding and using patient experience data within and alongside clinical trials.

EDITOR: Dr. Matthew Reaney

This file contains only Chapter 4 of this book. For a full copy of the book please email PCSBD@iqvia.com

Please cite this chapter as: Meyers O, Guillemin I, Reaney M (2023). Collecting patient experience data (PED) with clinical outcome assessments (COAs). In Reaney M (ed.) Using Patient Experience Data to Evaluate Medical Interventions. Generating, understanding and using patient experience data within and alongside clinical trials. IQVIA. Pages 43-56.

IQVIA

Table of Contents

Chapter 4

Collecting patient experience data (PED) with clinical outcome assessments (COAs)

O MEYERS, I GUILLEMIN, M REANEY

Key takeaways

- Clinical outcome assessments (COAs) are the most common method of generating patient experience data (PED) during intervention development. Best practices, evidentiary standards and a roadmap for developing "fit-for-purpose" COAs exist, but their application and consideration are often perceived lengthy and cumbersome.
- At the same time, strategies for capturing patient-relevant endpoints in clinical trials are constantly evolving. Forward-thinking research can contribute to intervention development and clinical care by the thoughtful collection and interpretation of COAs.

Advice for researchers interested in collecting PED with COA

- A clear definition/view of the final objective ahead of the COA measurement strategy design and implementation is essential: start with the end in mind.
- The use of a rigorous methodology, based on well-established principles yet flexible to the needs of the development or treatment setting, is essential to arriving at a COA measurement strategy that collects meaningful and interpretable PED.
- When developing a COA intended for labeling purposes, early interaction with the regulatory bodies is recommended in shaping a COA measurement strategy.

"Get your facts first, and then you can distort them as much as you please."

— *Mark Twain*

A clinical outcome assessment (COA) is a measure that describes or reflects how a patient feels, functions or survives.[1] COAs can provide data on patients' lived experience of a disease or condition, and a treatment or other intervention. They offer a direct, straightforward and systematic method for collecting patient experience data (PED) in the context of a clinical trial or clinical healthcare delivery. As such, COAs are widely used in both routine healthcare practice and intervention development settings.

The use of COAs to capture PED in intervention development is promoted by regulators in the U.S. and Europe. The U.S. Food and Drug Administration (FDA) has been at the forefront of establishing guidance for intervention developers and sponsors to encourage the proper collection and reporting of COA data, from the original 2009 patient-reported outcome (PRO) guidance[2] to a recent series of patient-focused drug development (PFDD) guidance.[1,3–5] Similarly, the European Medicines Agency (EMA) has long made consideration of the patient perspective a priority, and in their recent strategy for 2025, declared the importance of "ensuring the patient voice is systematically incorporated throughout drug development & associated evidence generation"[6] (see **Chapter 10** for more information on regulatory guidance on COA). These initiatives are bearing fruit in terms of the research practices of intervention developers/sponsors as well. For example, a recent review found that for the 5-year period from 2014–2018, 41 of 55 (75%) oncology drugs approved by FDA included PRO data collection in registrational trials.[7]

This chapter will focus on the science and strategy behind collecting PED with COA, including current best practice and future trends. It will describe the process for developing a COA measurement strategy for intervention development and outline the utility and challenges of using COA data to define benefit of treatment, as well as summarizing opportunities for improvement and innovation.

Why should I use a clinical outcome assessment (COA)?

A wide variety of medical technologies and tools, such as imaging, biopsy and laboratory tests, are available to healthcare professionals and researchers to assess physical and biological data. However, many health-related experiences, such as symptoms (like pain intensity or urinary frequency), daily impacts, treatment and healthcare management satisfaction, or treatment tolerability, can only be understood by collecting data directly from the patients themselves and/or people who know them well. A COA is a measure that describes or reflects how a patient feels, functions or survives[1] and offers a systematic approach to measure both observed and unobserved variables, such as perceptions or feelings. A COA can be used to establish the benefits and harms of an intervention under development from the perspective of the patient, caregiver or medical professional. It can also be used to engage patients in their healthcare management, and inform shared decision-making (see **Chapter 12** for more information on shared decision-making).

However, there are a lot of COAs to choose from, and new ones are constantly being developed, which sometimes renders the selection of the COA difficult. A solid and well-informed COA measurement strategy is therefore required, starting with the concept of interest, and ending with the implementation of the COA measure. This offers the greatest opportunity to comprehensively understand the benefits and harms associated with an intervention during its development.

How do I develop a COA measurement strategy?

There are numerous steps in selecting a COA to establish the benefits and harms of an intervention under development, including defining a concept of interest (COI), describing the context of use for the COA, deciding on a type of COA and examining whether it is

"fit-for-purpose." How the COA will be implemented also needs to be considered. These are outlined in **Figure 1** and described in detail below. A case study following this approach is thereafter outlined in **Figure 4**.[8]

Figure 1: Steps in Selecting a COA to Establish the Benefits and Harms of an Intervention Under Development

1. Defining the concept(s) of interest (COI) – what concept(s) do you want to measure, and why?
2. Describing the context of use – in what setting do you want to measure the concept(s)?
3. Deciding on the type(s) of COA – from whose perspective are the concepts best measured? Should a generic or context-specific measure be used?
4. Ensuring that COA measures are "fit-for-purpose" – are COA measures available to measure the COI in the context of use in a way that is well-defined, reliable, valid and interpretable?
5. Considering COA implementation – how will you implement the COA in the context of use? Is the full COA needed?

DEFINING THE CONCEPT(S) OF INTEREST (COI)

A COA is only as good as the outcome it measures. To this regard, PED research is essential to ensure an understanding of the disease, and identification of the COIs that could be used to determine the benefits and harm associated with an intervention. This is fundamental to an outcome strategy – you should first explore experiences with disease and treatment and determine which are the measurement concepts of greatest relevance to clinical decision-making, and which are most bothersome/impactful to patients. These should then be prioritized to establish the benefits and harms of an intervention under development. For example, is pain reduction a priority for patients? Would improving physical functioning be a meaningful outcome for patients?

Ideally, multiple sources of data should be used for to explore experiences with disease and treatment and define the COI.[9] Acceptable methods for defining the COI include a review of scientific peer-reviewed literature and other publications,[10] and discussions with key experts, including medical professionals, regulatory and payer stakeholders.[3] But qualitative "concept elicitation" work with patients is most important, as described in **Chapter 3**. A curated database that seamlessly indexes multiple sources of concept elicitation data, such as the IQVIA COA Accelerator (https://bit.ly/COAAccelerator), can facilitate decision-making around the COI.

The COI will be the basis from which a measurement is selected to determine the benefits and harm associated with an intervention. It is important to note at this point that the selected measure may be a COA, a biomarker, or a clinical exam – whichever is most appropriate to measure the COI. The important aspect at this stage is to understand what the patient believes is most important, in order to select the most appropriate measurement strategy.

DESCRIBING THE CONTEXT OF USE

The measurement strategy for the COI must be defined within a specific context of use. Elements that define the context of use include specific details about the target population, the disease, the study design, and the type of measurement planned to be used to determine the benefits and harm associated with an intervention.[1] If the researcher determines that a COI is best measured by a COA, they should then decide on the type of COA to use.

DECIDING ON THE TYPE(S) OF COA

There are four main types of COAs used in clinical research and routine healthcare: patient-reported outcome (PROs), observer-reported outcome (ObsROs), clinician-reported outcome (ClinROs), and performance outcome (PerfOs) (see **Figure 2**). All four COA types are evaluated using "measures" – with standardized measures often being used for PROs, ClinROs and ObsROs, and standardized tests being used for PerfOs.

PRO instruments capture data about the status of a patient's health condition or health-related topic directly from the patient, without interpretation of the patient's response by a clinician or anyone else.[11] PROs are often evaluated using questionnaires, diaries, and

related question-and-answer type formats – called PRO measures (PROMs) – and can be collected in paper format or using electronic data capture (e.g., application on a mobile device, web-based, or an interactive voice-response system [IVRS] conducted via telephone). PROMs are well placed to assess symptoms or other unobservable concepts that are only known by the patient, such as symptom severity, patient perception of their health status, perceived level of functional impairment, disability and health-related quality of life (HRQoL). PROMs are often administered alongside patient-reported experience measures (PREMs), with the latter more specifically reporting on patients' judgements of their experience of healthcare treatment and management while they are receiving care, rather than on the outcomes *per se*.

ObsRO instruments capture information about a patient's experience based on a report of observable signs, events or behaviors related to a patient's health condition by someone other than a patient or a health professional. They are best suited in instances where the patient is unable to reliably report for themselves, and are thus commonly used to evaluate outcomes in young children or individuals with cognitive impairment (see **Chapter 7** for more information on these populations). In these cases, caregivers who spend time with the patient in daily life can report on signs or behaviors they observe. ObsRO measures (normally questionnaires or diaries) can be collected in paper format or using electronic data capture (e.g., mobile phones, computers), capturing observations from a parent, a caregiver or someone who lives with or can observe the patient in their daily life. An ObsRO measure does not include medical judgment or interpretation.[11]

ClinRO instruments capture information about a patient's experience based on a report of observable signs or behaviors, like an ObsRO measure, but from the perspective of a trained healthcare professional, rather than a caregiver or similar. This allows for clinical events, or other manifestations related to a disease or condition that would benefit from clinical judgment – for example, the severity of a rash – to be captured.[11] ObsRO measures are common in Central Nervous System disorders (see **Chapter 7**). As with PROMs and ObsRO measures, ClinRO measures can be evaluated using questionnaires in paper or electronic format, but there is an increasing use of digital health technologies (DHTs) to support ClinRO evaluation. This may include use of photographs or recordings to evaluate a sign such as papulation or cough.

PerfO measures capture patient's experience based on a standardized task performed by the patient.[11] They are generally administered (with or without specific instructions) by a trained individual or completed by the patient independently. PerfO measures may be used to understand ability to perform cognitive or physical tests under standardized conditions. DHTs can also be used to collect patient performances through computing platforms, connectivity, software, and wearable sensors like smartwatches. For example, DHTs can be used to assess walking bouts in Parkinson's disease, breathing in COPD, and waking episodes in insomnia. DHTs are discussed in detail in **Chapter 6**.

COAs are further categorized based on the disease and population they are intended to. Generic COAs, where the items of the measure are adapted to any disease population and to the general population should be distinguished from disease-specific COAs, where the items are specific to a diagnostic group, a disease or a patient population. Disease-specific COAs are sometimes deemed more sensitive to change than generic measures because they measure aspects that are particularly salient to a specific disease or patient group. Generic COAs, on their side, should be considered when comparing between studies and diseases.

ENSURING THAT COA MEASURES ARE "FIT-FOR-PURPOSE"

When the COI is known, the context of use determined, and the specified COA type decided, the next step

Figure 2: Types of Clinical Outcome Assessments (COAs)

Patient Reported Outcome (PRO)
A measurement that comes *directly from the patient.*

Clinician Reported Outcome (ClinRO)
An assessment *determined by a trained medical professional.*

Performance Rated Outcome (PerfO)
A measurement based on a *task(s) performed by a patient.*

Observer Reported Outcome (ObsRO)
An assessment *determined by an observer* (i.e., a non-clinician, such as a teacher or caregiver).

is the select or develop a COA measure to establish the benefits and harms of an intervention under development. It is important to ensure that the COA measure is "fit-for-purpose"; that is it evaluates relevant and meaningful concepts and generates reliable, valid and interpretable data.[1] Ensuring that a COA measure is "fit-for-purpose" is a multi-step process involving both qualitative and quantitative research.[1,2]

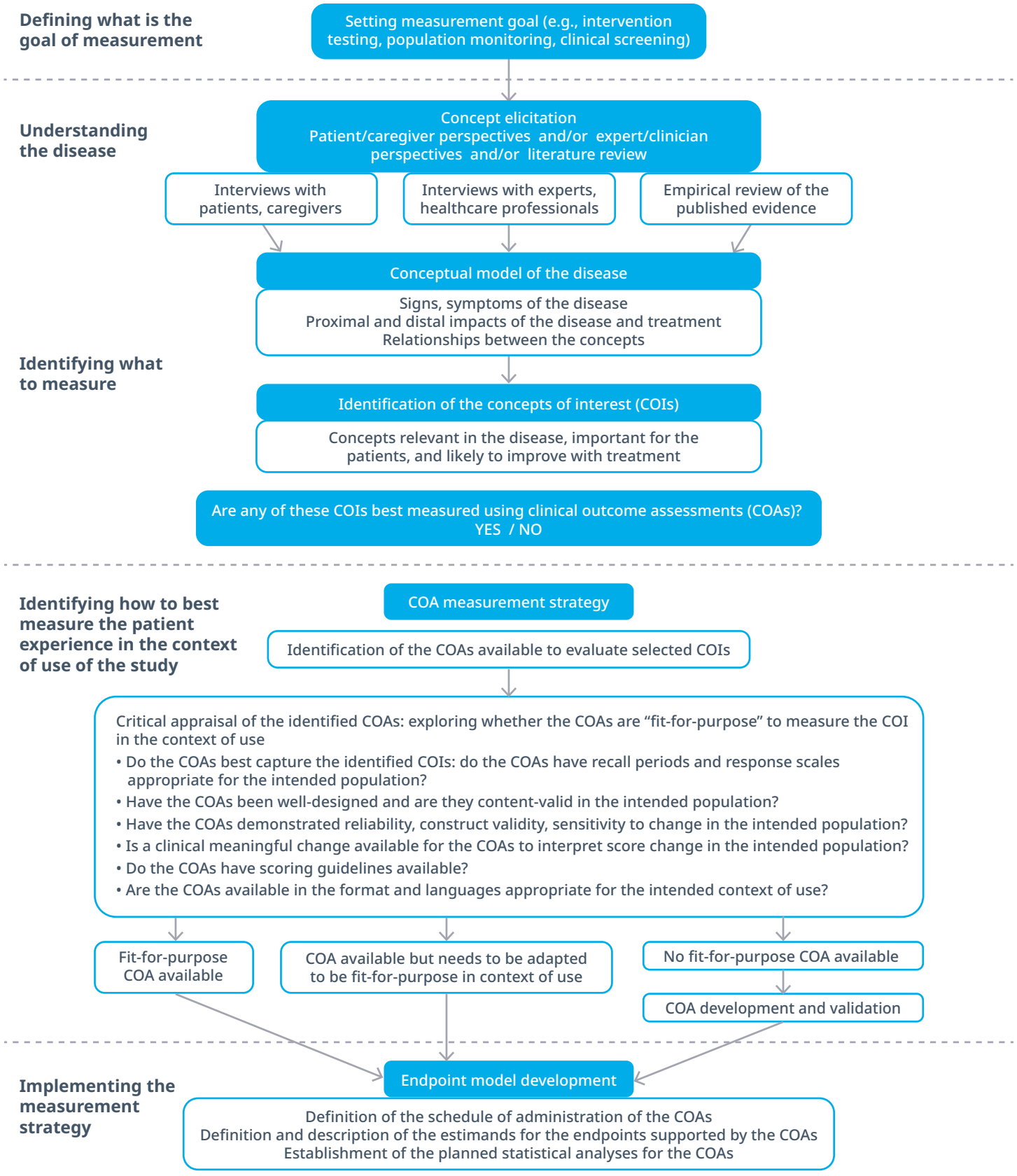
Qualitative "concept elicitation" research is essential to ensure that the COA measure is developed based on a robust understanding of the disease, and that it measures the concepts that are of greatest relevance to clinical decision-making, and which are most bothersome/impactful to patients (i.e., the COI).[12] Qualitative "cognitive debriefing" research is also essential to ensure that the language of the instructions, items, and response options in the COA measure is clear and comprehensible to every person that will be using them in the context of use (see **Chapter 3**).[13] This is collectively known as "content validity" – the ability of the COA measure to accurately assess aspects of the patient experience that are relevant and important and constitute meaningful aspects of health. Once content validity for a COA measure has been established, quantitative research is required to assess its psychometric performance (reliability, construct validity, sensitivity to change), and generate estimates of thresholds for defining treatment effects (meaningful changes) to support interpretability. This is described in detail in **Chapter 8**. A COA measure that is content-valid, psychometrically supported and interpretable could be considered as "fit for purpose." However, additional steps have sometimes to be considered depending on the study setting, like in the case of international clinical trials, necessitating culturally and conceptually equivalent (i.e., linguistically validated) COAs for the data to be compared and pooled.[14,15] Indeed, a COA may be "fit-for-purpose" in one context of use, but it cannot be assumed that it would also be considered "fit-for-purpose" in another.

A COA may be "fit-for-purpose" in one context of use, but it cannot be assumed that it would also be considered "fit-for-purpose" in another.

If an existing COA measure is available, in the public domain or through a licensing agreement, and can be considered "fit for purpose," this can be a very resource-efficient way to measure the COI. There are other instances when an existing COA measure is comprehensive but needs some modifications to be considered a comprehensive measure of the COI in the context of use. On other occasions still, if a thorough search of the published literature and other resources fails to turn up a "fit-for-purpose" COA measure, a new one may need to be developed. The process of developing a "fit-for-purpose" COA measure is shown in **Figure 3**. It begins with defining the COI, as described above. Next, a draft version of the COA is generated to measure that COI. When developing the draft of the COA measure, it is important to precisely articulate the COI to be measured and the target population and context of use, then define the item pool and the response scale (binary, ordinal, continuous and/or nominal; unipolar or bipolar). There is no rule regarding the number of items, but it is good practice to have as many as needed to capture the COI, albeit not too many. When a COA measure has been drafted, it should be tested through "cognitive debriefing," as described above to establish content validity. For a new COA measure this involves testing the item pool, along with the response scale and instructions and refining the contents based on feedback. In the case of PRO and ObsRO instruments, a wide range of reading ability and medical literacy must be taken into consideration when conducting cognitive debriefing research, especially when the respondents may include children and people with cognitive limitations.[16] Indeed, there is increasing interest in reducing or eliminating text from PRO instruments entirely in favor of graphical and other multimedia formats, in order to address concerns about patient literacy and comprehension.[17] Once content validity for a new COA measure has been established, its psychometric performance should be explored, as outlined above and detailed in **Chapter 8**.

The approach described above and in **Figure 3** is grounded in a roadmap produced by the FDA to describe how to develop a "fit-for-purpose" COA[1,2] which has since been adopted and referenced by other stakeholders interested in COA data – including the EMA, and multiple health technology assessment (HTA) and payer agencies. These stakeholders are keen to ensure that relevant COA data are both submitted for review, and that the data is well-defined and "fit-for-purpose." Regulatory and HTA/payer considerations for COA data are further outlined in **Chapter 10** of this book. However, while the approach outlined above undeniably has advantages and has proven its utility, it must be acknowledged that it is not always easy and sometimes not feasible to develop a COA measure as described. For example, in the case of rare diseases or with a specific age range where the access to population of interest may be limited, sample sizes for the quantitative research may be smaller than optimal. This is discussed in **Chapter 7**.

Figure 3: Steps in the Development and Implementation of a COA Measure (Adapted from Reaney et al.[18])



CONSIDERING COA IMPLEMENTATION

Once the decision is made to use COAs to establish the benefits and harms of an intervention under development, it is worthwhile to devise a strategy for implementation that properly balances the values of collecting complete and accurate data on the one hand and honoring the patient's experience of the process on the other. Whether the patient or caregiver is directly reporting (PRO and ObsRO), or is present at a study visit in order for data to be collected by a clinician (ClinRO) or via the performance of a task (PerfO), the quality and completeness of the data depend on cooperation from these stakeholders. Strategies to improve compliance and reduce missingness are an important consideration.

When using COAs in any context, be it a clinical trial or a clinical care setting, it is worthwhile to pay close attention to demands placed upon those providing the data. An evaluation of the burden to the COA respondent, which could be a patient, clinician or a non-clinician observer, would include noting the length of the COA in terms of number of items and time required to respond to them, as well as the complexity and cognitive load associated with completing the COAs and/or interacting with the measurement devices. Most PRO, ObsRO and ClinRO instruments or diaries consist of a set of instructions for the respondent and a series of prompts that elicit a response from a set of provided response options. The choice of format, or mode of administration (paper or electronic, should account for any known constraints in the user's ability to interact with one mode over another (considering, for example, sensory and graphomotor abilities) and the suitability of the COA for a particular mode of administration (e.g., a visual analogue scale that may be better administered on paper/screen than by phone). There is an increasing use of electronic formats for administering COAs because of the many advantages they display over their paper format counterpart; those are detailed in **Chapter 6**. This is particularly true for clinical trials. Electronic COAs (eCOAs) correspond to the paper-based version of a COA that has been migrated onto an electronic format. As such, development and validation should follow the same standardized methodology as the one described herein. However, to ensure of the usability and functionality of this electronic format, an additional validation step is recommended, namely usability testing. Its purpose is to assess the abilities of the population to use the measure easily, to understand its instructions and follow them accurately, and to evaluate the functional aspects of the eCOA, such as the format, screen size and color, and font size (see **Chapter 6**). Patient and caregiver input in the design process can further help to consider the number of COA items that can be tolerated at a visit or within a diary, the relevance of question topics to their experiences, and even the formatting of measures that can impact not just the content validity and quality of the data, but even the likelihood of respondents correctly and willingly providing data throughout the course of a trial or treatment course (see **Chapter 9** for more discussion on patient engagement to facilitate completion of COA measures).

Implementing COAs in the context of a clinical trial requires the additional consideration of several study-specific questions. In particular, the schedule of administration requires careful thought. Variables to consider include study parameters, such as the efficacy endpoint and the length of the measurement period, as well as instrument parameters, such as the recall period. Depending on the therapeutic area, the endpoints to be measured, and the trial design, it may be more appropriate to conduct COA measurement at study visits or send patients home with an electronic diary (or a hybrid approach). With the increasing attention being paid towards the adoption of decentralized clinical trials and remote decentralized clinical trials, novel approaches to scheduling assessments, including COAs, are sure to follow.

Figure 4: Case Study of a New COA Measure Developed to Establish the Benefits and Harms of an Intervention Under Development

BACKGROUND AND OBJECTIVE

Progressive familial intrahepatic cholestasis (PFIC) is a rare disease affecting infants and young children. Invasive surgery is the only approved treatment option. An intervention developer/sponsor was developing a new treatment, odevixibat (Bylvay). They required a clinical outcome assessment (COA) to support the primary efficacy endpoint (focused on core symptoms) in the pivotal Phase III clinical trial.

CLINICAL OUTCOME ASSESSMENT MEASUREMENT STRATEGY[19]

- A literature review was conducted and supplemented with clinician interviews, and concept elicitation interviews with caregivers and patients to identify the core symptoms relevant to patients/caregivers and which may be an appropriate marker of treatment benefit for odevixibat.
  - » The research identified pruritus as a symptom that has a huge impact on children with PFIC. It is a deep, relentless itch which is particularly strong at night. Children scratch themselves until they bleed.
  - » The sleep disruption and the pain from their damaged skin affects their social and educational development, and the quality of life for them and their families
- A novel, content-valid patient-reported outcome (PRO) daily diary was developed to measure itch with the help of the wording used by patients/caregivers during the concept elicitation interviews.
  - » The majority of patients in the pivotal trial were expected to be <8 years, therefore the measure that was developed combined an ObsRO (observer reported outcome) and a PRO (for the older children)
  - » Cognitive debriefing interviews with patients and caregivers were conducted to confirm the newly developed ObsRO and PRO daily diary were understandable and interpreted consistently and as intended by the caregivers and children with PFIC, respectively
  - » The new measures were translated and culturally adapted for the Phase III trial according to best practices
- A patient-focused endpoint focused on pruritus was defined using the new measures.
- The psychometric properties of the ObsRO and PRO daily diary were evaluated within the clinical trial dataset.
  - » The analyses confirmed that the diary was clinically valid, reliable and sensitive to change[20]
  - » An estimate of a meaningful change threshold was generated

EMA AND FDA REGISTRATION APPROVAL AND DISSEMINATION TO THE SCIENTIFIC COMMUNITY

- The Phase III trial met both primary endpoints for odevixibat in PFIC. Both EMA and FDA approved odevixibat as the first drug to treat PFIC.
- A comprehensive PRO evidence dossier was submitted alongside the registration filing. The pruritus-related endpoint was included in product labeling.
- Publications were developed for the ObsRO/PRO diary development to share the data with the scientific community.[19,20]

What can I do with COA data?

COA data can be used to support regulatory, HTA and clinical decision-making.

Although there are numerous ways to generate PED (see **Chapters 3, 5 and 6** of this book), COA currently provides the most valuable PED for regulators in benefit-risk analyses and approval decisions. This is due to COAs being most likely to support endpoints in clinical trials, with other types of PED more often used for background and context to understand patient priorities and preferences, the disease or condition, or burden on patients or caregivers. COA endpoint data may be included in regulatory product labeling if derived from fit-for-purpose COA measures implemented in well-developed clinical trials. COA data can also be used in HTA assessment; indeed, they are the only PED that fit into the formal assessment framework of HTAs. See **Chapter 10** for a summary of how COA has been used in regulatory and HTA decision-making.

COA data is also increasingly important in "shared decision-making," where clinicians and patients come together to make decisions about care and treatment. Both patients and clinicians find COA data useful to better understand how a treatment can impact – positively or negatively – on feelings and functioning. For example, a greater understanding of the tolerability of treatment, collected through COA, can help clinicians set and manage expectations with patients should that treatment be selected for use. See **Chapters 11 and 12** for a summary of how COA can be shared with medical and patient communities to facilitate its use.

What will happen with COAs in the future?

As described above, traditional COAs (PRO, ClinRO, and many ObsRO instruments) take the form of instructions and a static set of items with associated response options. The time-honored tradition in COA use, taken from its roots in psychological testing, was fidelity to the measure: any change in administration, especially a change to the sequence of the items or the use of anything less than the full set of items, was considered to be a threat to the reliability and validity of the measure, calling into question any interpretation that might be made of the resulting data. For several different interlocking reasons, over the last number of years, this attitude has begun to shift. These reasons include the advent of test construction theories other than classical test theory, and the desire to reduce the length of measures and thus the burden to respondents. There is an increasing use of banks or libraries of items in the course of the development of new scales and/or the generation of customized assessments in real time. There are several different manifestations of this move to acceptance of non-static implementation of COAs.

One manifestation is item libraries. Some item libraries comprise items from measure that may have been developed decades in the past. For example, both the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaires (EORTC-QLQ)[21] and Functional Assessment of Chronic Illness Therapy (FACIT)/Functional Assessment of Cancer Therapy (FACT)[22] families of questionnaires have procedures by which licensed users can check items out of the library and compose bespoke questionnaires. Other item libraries are sets of items that were developed, at their outset, to be used selectively. Prominent examples are the PRO version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE),[23] which has more than 120 items covering nearly 80 different measurement concepts, and the IQVIA Covid Daily Diary (ICDD),[24] which measures symptoms and impacts of COVID-19 and impacts of the pandemic across 88 items. Users generally select a subset of the items for a particular use case.

Sets of items detected from item libraries which are believed to measure a single underlying construct (symptom, functional area, meaningful aspect of

health, etc.) are known as item banks. The items in the item bank are subjected to statistical analysis, using approaches such as item response theory (IRT) and Rasch modelling (see **Chapter 8** for more information) intended to result in one or more new measures which theoretically capture the appropriate range of clinical severity of the population of interest. The Patient-Reported Outcomes Measurement Information System (PROMIS) is an example of this approach.[25] Taking the item banking concept one step further, computer adaptive testing (CAT) uses information about the items in a bank to determine the next item administered to a respondent in real time, based on their previous responses. An example of a measurement system that includes a set of instruments based on CAT is Quality of Life in Neurological Disorders (Neuro-QoL),[26] a close cousin to PROMIS.

Apart from selecting customized sets of items is the notion that specific aspects of the patient experience can be measured and tracked, personalized to each patient's report of the symptom or symptoms that are the most severe, most bothersome, or which they are most keen to see improvement in. The idea of personalized endpoints has allure for the ability of intervention developers and sponsors to show within-patient response to treatment based on the aspects of the disease experience that are most relevant to each individual patient.[27] Even so, there remain significant administrative and analytic challenges and barriers to acceptance at the national stakeholder level (i.e., regulators and payers).

Other ways that the science of COAs is developing and adapting to the times includes more robust approaches to the use of digital endpoints (refers to **Chapter 6** for detailed discussion). According to the Digital Medicine Society (DiMe), as of this writing, more than 550 users have interacted with their Sensor Data Flow Design Tool, 101 intervention developers and sponsors have collected digital endpoints in clinical studies, and more than 366 unique digital endpoints have been implemented.[28]

Collecting PED with COAs: A summary

The development and implementation of a well-considered COA measurement strategy as part of a PED initiative in medical product development or in clinical care to establish the benefits and harms of an intervention is an extensive and lengthy process, requiring qualitative and quantitative research to ensure that the right concept(s) are measured using "fit-for-purpose" COAs, and that they are implemented in a way that maximizes data collection and aids interpretability. Given the costs in time and money associated with it, a fair question to ask is, why consider making this investment?

COAs help researchers tell a detailed, evidence-based story about the patient and the benefits and harms of new treatments, interventions or management. Robust implementation of a well-considered COA collection strategy in medical product development can lead to tangible benefits for intervention developers and sponsors. First and foremost, when COA data collected in the context of a well-designed study are illuminating as to the efficacy and in some cases, safety with the respect to the use of a particular medical product, this can be negotiated by the sponsor into the medical product label/package insert/summary of product characteristics. In other words, a regulatory label claim. In the United States in particular, this translates into the restricted permissibility to use PED in direct-to-patient marketing, and generally this means that there is regulatory endorsement for using the PED as a component of commercial promotional activities. Further, certain COA data will be considered by HTAs in setting the reimbursement level, and the data can be strategically communicated to healthcare professionals and patients and caregivers about the likely experience of patients in the clinical setting (see **Chapters 10–12**). In clinical care context, engaging patients through the implementation of "fit-for-purpose" COAs helps in understanding

patients' needs and experience, and in turn contributes to shared decision-making. Engaging and empowering patients in their healthcare is key in their adoption of this management and paves the road to success.[29]

The collection of PED through COAs is a mature field whose roots go back decades or more, although due to innovations in science and technology, it remains a young field with much promise yet to be realized. Ongoing efforts are seeking the establishment of minimum COA measure sets in particular disease areas. This important work can help provide a level of standardization and comparability, and is already beginning to bear fruit, but requires additional time and effort to mature. Additional areas for maturation of the field are the definition of best practices for COA collection in special populations (see **Chapter 7**), and an understanding of the methods and best practices for employing a completely individualized approach to COA.

Altogether, the implementation of a "fit-for-purpose" COA measurement strategy is an extensive and lengthy process, and therefore it is preferable that this strategy takes place as early as possible in the clinical development program, ideally prior to the implementation of COA in the pivotal trial, and certainly with early interactions with the targeted regulatory authority(ies) to endorse this strategy.

References

1 U.S. Food and Drug Administration (FDA). Patient-Focused Drug Development: Selecting, Developing, or Modifying Fit-for-Purpose Clinical Outcome Assessments (PFDD Guidance 3 [draft]); 2022. https://www.fda.gov/media/159500/download. Accessed April 4, 2023.

2 U.S. Food and Drug Administration (FDA). Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. Guidance for Industry. December 2009. https://www.fda.gov/media/77832/download. Accessed April 12, 2023.

3 U.S. Food and Drug Administration (FDA). Patient-Focused Drug Development: Collecting Comprehensive and Representative Input (PFDD Guidance 1 [final]); 2020. https://www.fda.gov/media/139088/download. Accessed April 4, 2023.

4 U.S. Food and Drug Administration (FDA). Patient-Focused Drug Development: Methods to Identify What Is Important to Patients (PFDD Guidance 2 [final]); 2022. https://www.fda.gov/media/131230/download. Accessed April 4, 2023.

5 U.S. Food and Drug Administration (FDA). Patient-Focused Drug Development: Incorporating Clinical Outcome Assessments Into Endpoints for Regulatory Decision-Making (PFDD Guidance 4 [draft]); 2023. https://www.fda.gov/media/166830/download. Accessed April 4, 2023.

6 European Medicines Agency. EMA Regulatory Science to 2025 Strategic reflection. https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/ema-regulatory-science-2025-strategic-reflection_en.pdf. Accessed January 16, 2023.

7 Gnanasakthy A, Gnanasakthy A, Russo J, Gnanasakthy K, Harris N, Castro C. A review of patient-reported outcome assessments in registration trials of FDA-approved new oncology drugs (2014–2018). *Contemp Clin Trials.* 2022;120:106860.

8 Ho CN, O'Quinn S, Bailey J, Meyers O, Slagle AF, Dellon ES, et al. Development and content validation of a symptom assessment for eosinophilic gastritis and eosinophilic gastroenteritis in adults and adolescents. *Orphanet J Rare Dis.* 2021;16(1):495.

9 Meyers O, Reaney M. PCN321 PRO conceptual models in lung cancer: The value of multiple data sources. *Value Health.* 2020;23(Supplement 2):S479.

10 Meyers O. 3151 Concept elicitation via structured review of memoirs: The ISROM Method. Presented at ISOQOL, October 2021, Virtual.

11 U. S. Food and Drug Administration (FDA)-NIH Biomarker Working Group. BEST (Biomarkers, EndpointS, and Other Tools) Resource. Published 2021, updated 2021. https://www.ncbi.nlm.nih.gov/books/NBK338448/. Accessed February 7, 2023.

12 Patrick DL, Burke LB, Gwaltney CJ, Leidy NK, Martin ML, Molsen E, et al. Content validity—establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: Part 1—eliciting concepts for a new PRO instrument. *Value Health.* 2011;14(8):967–77.

13 Patrick DL, Burke LB, Gwaltney CJ, Leidy NK, Martin ML, Molsen E, et al. Content validity—establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO Good Research Practices Task Force report: Part 2—assessing respondent understanding. *Value Health.* 2011;14(8):978–8.

14 Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, et al. Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value Health.* 2005;8(2):94–104.

15 Eremenco S, Pease S, Mann S, Berry P, and on behalf of the PRO Consortium's Process Subcommittee. Patient-Reported Outcome (PRO) Consortium translation process: consensus development of updated best practices. *J Patient Rep Outcomes.* 2017;2(1):12.

16 Ruzich E, Meyers O. PMU60 Assessing the readability of patient-reported outcome measures in trials for drugs with recent FDA approval. *Value Health.* 2021;24(Supplement 1):S155.

17 Long C, Beres LK, Wu AW, Giladi AM. Developing a protocol for adapting multimedia patient-reported outcomes measures for low literacy patients. *PLoS One*. 2021;16(6):e0252684.

18 Reaney M, Black P, Gwaltney C. A Systematic Method for Selecting Patient-Reported Outcome Measures in Diabetes Research. *Diabetes Spectr.* 2014;27(4):229–32.

19 Gwaltney C, Bean S, Venerus M, et al. Development of the Patient- and Observer-Reported PRUCISION Instruments to Assess Pruritus and Sleep Disturbance in Pediatric Patients with Cholestatic Liver Diseases. *Adv Ther.* 2022;39:5126–43.

20 Gwaltney C, Ivanescu C, Karlsson L, et al. Validation of the PRUCISION Instruments in Pediatric Patients with Progressive Familial Intrahepatic Cholestasis. *Adv Ther.* 2022;39:5105–25.

21 EORTC Quality of Life. Quality of Life Group Website. https://qol.eortc.org/. Accessed February 27, 2023.

22 FACIT Group. https://www.facit.org/. Accessed February 27, 2023.

23 NIH. National Cancer Institute. Division of Cancer Control and Population Sciences. Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE®). https://healthcaredelivery.cancer.gov/pro-ctcae/. Accessed February 27, 2023.

24 Reaney M, Turnbull J, Paty J, Heuer K, Gwaltney C. Development of an Item Bank to Assess Patient-Reported Outcomes: Signs, Symptoms, and Impacts of COVID-19. *Patient*. 2022;15(6):703–13.

25 NIH. National Institute of Aging. Patient-Reported Outcomes Measurement Information System (PROMIS). https://www.nia.nih.gov/research/resource/patient-reported-outcomes-measurement-information-system-promis. Accessed February 27, 2023.

26 Gershon RC, Lai JS, Bode R, Choi S, Moy C, Bleck T, et al. Neuro-QOL: quality of life item banks for adults with neurological disorders: item development and calibrations based upon clinical and general population testing. *Qual Life Res.* 2012;21(3):475–86.

27 Iwashyna TJ, Deane AM. Individualizing endpoints in randomized clinical trials to better inform individual patient care: the TARGET proposal. *Crit Care.* 2016;20:218.

28 Digital Medicine Society (DiMe). Advancing digital medicine to optimize human health. https://dimesociety.org. Accessed February 27, 2023.

29 Hickmann E, Richter P, Schlieter H. All together now – patient engagement, patient empowerment. *BMC Health Serv Res.* 2022;22(1):1116.