

**HTAi Annual Meeting: The Road to Policy and Clinical Integration
Adelaide, Australia: 26 June 2023
Panel Session Report**

Rare Disorders - Is the Lack of Effect in a Patient-Reported Outcome, Reflecting no Benefit?

Summary

- Many rare diseases are systemic, complex conditions that affect several organs often progressively over a lifetime. This leads to heterogeneous disease presentation, progression and response to treatment, even in one phenotypic or genotypic group. Thus in evidence generation, and HTA, it may be unclear in the early stages of documenting the natural history of a disease which endpoints should be used to prove added patient benefit of a new treatment.
- Given the spectrum of severities of most rare diseases, measuring quality of life is important, both in terms of impacts of the condition, and the impacts of treatments.
- A range of approaches may be undertaken to capture how a rare condition affects a patient, their family and other carers, including disease specific Patient-Reported Outcomes Measures (PROMs), surveys, questionnaires etc.
- Measuring quality of life is difficult, and even in prevalent conditions PROMs may lack sensitivity, be subject to more missing values than clinically reported outcomes and be analysed in underpowered studies. In small, heterogeneous rare disease populations, this is even more challenging.
- PROMs also add a burden of data collection on people living with the condition.
- Administration of PROMs may be difficult or impossible in some rare disease populations if they occur in babies/infants or result in cognitive or speech deficits, so development of observer reported outcomes may be needed.
- Further research is needed to develop robust, specific measures of the impact of a specific condition (and a treatment) on carers and families.
- Activities to understand the burden of disease should be undertaken in a multi-stakeholder manner, across Health Technology Developers. Rigorous methods should be used, and findings published in peer review journals to enable knowledge development across the rare disease community.
- Digital technologies and artificial intelligence offer new, reliable ways for continuously capturing data that can demonstrate impacts on quality of life, e.g. through use of wearables or real time image processing from video recordings.
- PRO evidence relating to treatment effects needs to take account of the challenges in developing and administering PROMs in the specific condition being studied and appropriate leniency applied, or greater uncertainty accepted. This requires an early, open dialogue between the Health Technology Developer and Health Technology assessor. The developer needs to explain their rationale for choice of endpoints and any challenges in administration and analysis, and the assessor needs to be open to discuss, accept, or even propose alternative and complementary sources of evidence.
- Other forms of evidence and input are needed to understand the quality of life impacts of treatment, such as robust patient-based evidence from qualitative research, data collected from clinicians and patients and systematic input of clinical experts and people living with the condition. New methods are needed to include expert generated data and expert opinions in the deliberative appraisal process.
- If a valid disease specific PROM can be developed, it could be used as part of a core dataset that is used to demonstrate value post launch.

Contents

Glossary	2
1. Rare Disorders: Setting the Scene from a Health Technology Developer's Perspective	3
2. Developing a PRO for a Rare Disease: The Experience of an International Patient Group ..	5
3. Recommendations for better use of quality of life evidence in HTA for rare diseases	7
4. A Clinician's Perspective on a New Paradigm for Evidence Generation of Patient Benefit	13
5. Discussion	18

Glossary

COA	Clinical Outcomes Assessment
EC	European Commission
EMA	European Medicines Agency
EU	European Union
GARDIAN	Gaucher Registry for Development, Innovation and Analysis of Neuronopathic
G-BA	Gemeinsamer Bundesausschuss
HAS	Haute Autorité de Sante
HSUV	Health State Utility Value
HTA	Health Technology Assessment
HTAi	Health Technology Assessment international
IGA	International Gaucher Alliance
NICE	National Institute for Health and Care Excellence
ObsRO	Observer Reported Outcome
OMP	Orphan Medicinal Product
PRO	Patient Reported Outcome
PROM	Patient Reported Outcome Measure
QALY	Quality Adjusted Life Year
RCT	Randomized Controlled Trial

About HTAi

Health Technology Assessment international (HTAi) is the global scientific and professional society for all those who produce, use, or encounter health technology assessment (HTA). As an open platform for global collaboration, HTAi leverages our shared and collective intelligence to advance and promote HTA worldwide. HTAi represents 66 organizations and over 800 individuals from more than 65 countries, including researchers, agencies, policy makers, industry, academia, health service providers, and patients/consumers. This panel was judged by three reviewers and selected for presentation at the 2023 annual meeting by the International Scientific Program Committee.

Status of this report

This report has been prepared by Karen Facey and approved by all presenters to become a public record. The statements attributed to panellists are their own opinion and not those of their organization or other panellists.

This report should be cited as:

Granados A, Collin-Histed T, Nicod E, Facey K, Iorio A. *Rare Disorders - Is the Lack of Effect in a Patient-Reported Outcome, Reflecting no Benefit?* Report of HTAi 2023 Panel Session. Health Technology Assessment International – Canada. 2023.

Funding

Karen Facey received a fee from Sanofi to act as rapporteur and support for travel to Australia.

Alicia Granados receives a salary from Sanofi.

Tanya Collin-Histed, Elena Nicod and Alfonso Iorio received expenses from Sanofi to support travel to the meeting.

1. Rare Disorders: Setting the Scene from a Health Technology Developer's Perspective

Alicia Granados MD PhD - Sanofi

Dr Granados gave an overview of the challenges generating evidence about the value of new medicines for rare diseases and the potential benefit of using Patient-Reported Outcomes (PROs), with a focus on genetic, metabolic chronic rare disorders. Research has identified 6,172 unique rare diseases¹ using the European Union (EU) orphan medicinal product (OMP) prevalence definition. This research also identified that for those rare conditions with recorded prevalence, 85% (4,508) are "very rare" with a prevalence of <1 in 1 million people, and 70% have exclusively paediatric onset.

In the case of ultra-rare, genetic, metabolic diseases, which are the focus of this panel session, it is not just their low prevalence that creates challenges in studying rare diseases, it is their complex pathophysiology. Most rare diseases are systemic (often affecting several organs leading to a range of clinical manifestations) and many are heterogeneous, with different sub-populations due to different genetic sub-types or other known or unknown aspects. This heterogeneity leads to different patterns of presentation of disease, progression of disease and response to treatment. This means the many rare diseases are complex in terms of characterization, evidence generation and evidence assessment.

Policies have been put in place to support development of treatments for rare diseases (such as the EU OMP regulation) and this has led to more regulatory approvals of treatments for rare diseases. However, some of these treatments have not been reimbursed due to perceived major uncertainties in clinical and/or cost effectiveness, often with a criticism that patient benefit in terms of improvements in quality of life have not been demonstrated.

CIRS (2022)² reported on the difference in HTA recommendations between 2017 and 2021 for OMPs in six countries in the EU and Australia, identifying that specific appraisal routes for rare diseases were not used by all countries and time for appraisal was generally longer for OMPs than more prevalent conditions.

Rare disease heterogeneity and the other issues associated with studying rare diseases make evidence generation difficult and this leads to uncertainties. Now all evidence has uncertainties but in rare diseases, these uncertainties should be evaluated in terms of what is feasible in that rare disease population, how meaningful the outcome is and whether the statistical analysis is capturing true value.

Health Technology Developers spend months consulting with a range of stakeholders in advance of the finalization of a protocol to agree the Clinical Outcomes Assessments (COAs) that should be included in their pivotal clinical studies. Many follow FDA guidance that considers the value of

- clinician-reported outcomes
- performance outcomes (standardized task administered by a trained individual)
- patient-reported outcomes
- observer-reported outcomes (signs, events or behaviours reported by someone other than the clinician or patient)³.

¹ Wakap et al. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6974615/pdf/41431_2019_Article_508.pdf

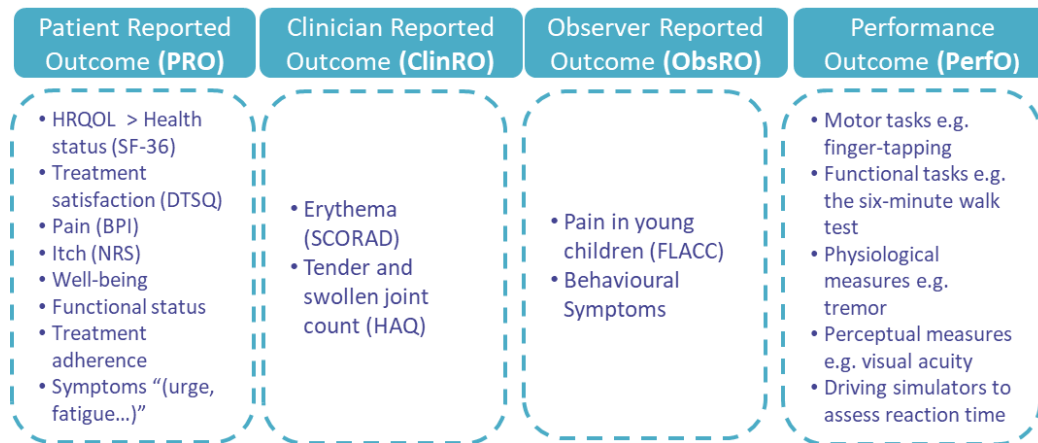
²Sola B, Wang T, McAuslane N. 2022. *R&D Briefing 86: Review of HTA outcomes and timelines in Australia, Canada and Europe 2017-2021*. Centre for Innovation in Regulatory Science. London, UK.

³ <https://www.fda.gov/about-fda/clinical-outcome-assessment-coa-frequently-asked-questions#COADefinition>

Examples of these outcomes are presented in Figure 1.

Figure 1. Different forms of Clinical Outcomes Assessments

Adapted from-<https://www.fda.gov/drugs/drug-development-tool-ddt-qualification-programs/clinical-outcome-assessment-coa-qualification-program>



COAs may be measured in clinical studies or in real-world settings, but to play a key role in regulatory or HTA/Payer decision-making they must be:

- ✓ well-defined
- ✓ cognitively appropriate
- ✓ consistently understood
- ✓ reliable
- ✓ valid
- ✓ generalizable
- ✓ sensitive to the disease severity/stage & treatment
- ✓ clear definition of meaningful change.

This HTAi panel focuses on the use of PROs and whether these pre-requisites for use are tenable in rare diseases. For example, in some rare diseases cognition is impaired, PRO measurement (PROMs) instruments may not have been validated for the specific rare condition or in that population (e.g. children), interpretation of meaningful change may not be agreed or take account of the heterogeneity of the disease. Then there are questions about how non-significant effect of a PRO should be interpreted. Does it mean no benefit for the patient, or is the PRO and its analysis inadequate to demonstrate a treatment effect?

The inadequacy of a PROM may arise because of the poor understanding of natural history, burden of illness, the small sample size and outliers due to the heterogeneity of disease. Furthermore in chronic diseases it is well known that patients learn to live with their disease and so a patient would score their quality of life higher than a member of the general population. These aspects lead to variability in the way in which PROs are interpreted in different health jurisdictions.

Concerted, multi-stakeholder, interdisciplinary approaches are needed for rare disease medical product development, evidence generation and appraisal. Patients and carers should be engaged early to enhance the ability to understand and integrate their unique and important perspectives to help advance and improve evidence generation.

PROs could be an important part of a more diverse evidence base and used to confirm other uncertain effects, but different methodological constructs and lenient assessment approaches are needed that take account of the feasibility of measurement and limitations of analysis.

The multi-stakeholder panel will discuss this further:

- Tanya Collin-Histed: CEO, International Gaucher Alliance
- Elena Nicod PhD: Director, Dolon Ltd
- Alfonso Iorio MD PhD: Chair Department of Health Research Methods, Evidence and Impact, McMaster University, Canada

2. Developing a PRO for a Rare Disease: The Experience of an International Patient Group

Tanya Collin-Histed - International Gaucher Alliance

Ms Collin-Histed gave an overview of how an international patient group developed a PRO and Observer Reported Outcome (ObsRO) to show the impacts of a complex, rare, disease with paediatric onset and its use in a global disease registry.

Gaucher Disease is a genetic disease in which fat-laden Gaucher cells build up in areas such as the spleen, liver and bone marrow, causing inflammation and dysfunction. Type II Gaucher Disease (typically diagnosed at 6 months of age) and type III (diagnosed at 6-18 months of age) additionally affect the central nervous system and brain (neuronopathic). Alongside the symptoms caused by the impacts on other organs, this leads to decreased physical functioning, and in some patients, recurrent seizures. Patients may require assistance for daily living (eating, washing, walking) and experience extreme fatigue. This may not only impact life expectancy but can also affect education, social interactions and mental health in terms of low self-esteem, depression, anxiety. As Dr Granados indicated, physical presentation and disease progression can be different among patients who are labelled as having similar disease. This is illustrated in Figure 2, which shows five patients with type III Gaucher disease and the same genotype who have very different clinical presentations. Whilst two of these patients have died, three are living fairly normal independent lives.

Figure 2: Neuronopathic Gaucher Disease Heterogeneity



This heterogeneity makes it challenging to study the disease. Hence although the first clinical trial of a treatment in this condition failed to show a benefit 20 years ago, there is still no good evidence about the natural history of the disease, no tool to measure neuronopathic impacts of the disease and no authorised treatments for neuronopathic aspects.

Hence the International Gaucher Alliance (IGA) decided to develop validated measures of the impacts of the disease on physical, cognitive, social and emotional functioning, in the form of a PRO and an ObsRO. These measures have been incorporated into a web-based platform to allow systematic data collection globally, with the aim of developing evidence that is sufficiently robust for use by Health Technology Developers and decision makers, such as HTA bodies.

The platform is the **Gaucher Registry for Development, Innovation & Analysis of Neuronopathic disease: GARDIAN**^{4,5}, which aims to develop knowledge about type II and III Gaucher Disease that will improve patients' lives. The objectives of GARDIAN are to identify where Gaucher Disease type II and III patients are, characterize their disease presentation and progression, and use of health and social care services. This should improve understanding among all stakeholders about the burden and impacts of the disease over time and help identification of patients for clinical trials and inform choice of endpoints that demonstrate patient benefit.

GARDIAN is global registry that is patient-owned, and patient led by the IGA. It began with a feasibility study in 2019 and was launched in April 2022 as a secure data collection platform. At the outset the registry is being made available in seven languages: English, Spanish, Arabic, Japanese, Chinese, German and French, to cover the largest portions of the global population.

Patients/carers upload their diagnostic information and clinical experts verify that they have the neuronopathic form of Gaucher Disease. Then patients/carers enter patient information onto the registry. A baseline questionnaire captures demographics, social and economic indicators, time of diagnosis, symptoms at diagnosis and at entry. Other data captured includes treatments, aids required for daily living and use of health and social care services.

Early data from the registry on 13 patients indicate that a third have been diagnosed with anxiety and depression, 23% have cardiac abnormalities and 15% have liver disease, while 23% use wheelchairs and 23% require institutional long-term care.

Regular data entry is encouraged every six months, including PROMs for issues known to be of importance in the condition, such as depression (PHQ-9), anxiety (GAD-7) and paediatric quality of life (PedsQL). In addition to these generic PROMs it was agreed that disease-specific quality of life measures should be developed in the form of a PRO, and as the condition has onset in babies, an ObsRO.

The PRO and ObsRO measurements were developed over three phases, in partnership with patients, carers and clinicians with expertise in treating neuronopathic Gaucher Disease. Phase 1 gathered qualitative input from a systematic literature review, consensus discussions with clinical experts, and focus groups and interviews with patients and carers to identify concepts for inclusion in the outcome measures that would capture aspects of the disease that mattered to them.

⁴ www.gardianregistry.org

⁵ Collin-Histed, T *et al.* A global neuronopathic gaucher disease registry (GARDIAN): a patient-led initiative. *Orphanet J Rare Dis* 2023; **18**: 195 <https://doi.org/10.1186/s13023-023-02828-w>

This included issues that are not traditionally studied in clinical trials such as pain, fatigue, emotional distress, disability, treatment concerns, experience of care, unmet needs etc. and resulted in domains that covered:

- Symptoms
- Physical function
- Self-regulation
- Social-emotional function
- Cognitive function.

This led to the inclusion of the existing relevant symptom and generic PROs into GARDIAN and development of a new disease specific PROM containing 33 questions covering:

- Burden of disease - symptoms
- Physical functioning – self-care, feelings
- Daily and social activities
- Health and travel concerns
- Overall health

In addition, there are six questions to assess the impact on carers.

Phase 2 of the development was cognitive interview testing of UK and US English versions of the draft outcome measures to ensure that the way in which the information had been captured in the questions was understandable to patients and carers. This led to changes in the wording of questions in the instrument.

Phase 3 then involved translations of the questionnaires into French, German, Spanish, Arabic, Japanese, Chinese and similar cognitive interviews to check clarity of wording.

The aim is to achieve psychometric validation of the PROM and ObsRO within GARDIAN so that they can be used to demonstrate the impacts of neuronopathic Gaucher Disease and be used in clinical trials in the future.

3. Recommendations for better use of quality of life evidence in HTA for rare diseases

Elena Nicod PhD - Dolon Ltd (and Karen Facey PhD – University of Edinburgh)

Dr Nicod gave an overview of research about the use of PROMs in HTA of rare diseases that had been undertaken in the EC-funded IMPACT HTA project⁶.

Quality of life impacts are an important measure of burden of disease and the consequences of treatment. However, demonstrating the impacts of treatments on quality of life is not easy, even in prevalent conditions. As a result, guidance on measurement of quality of life has been developed in recent years, most notably by the FDA in terms of how to develop robust patient-focused outcome measures, including PROMs⁷.

⁶ https://www.impact-hta.eu/_files/ugd/e1a359_4c6f61bc3cc94ae6a7781490d8b7ab49.pdf?index=true

⁷ <https://www.fda.gov/drugs/drug-development-tool-ddt-qualification-programs/roadmap-patient-focused-outcome-measurement-clinical-trials-text-version>

PROMs can be classified as

- Generic: measuring general elements of quality of life such as physical and social functioning, physical and emotional limitations, mental health, pain etc. These can be used in any disease and compared across diseases (e.g. EQ5D⁸, SF36⁹).
- Disease-specific: recording impacts identified as important in specific conditions (e.g. DMD-QOL¹⁰). These are intended to be more sensitive than generic PROMs but cannot be used to compare across diseases.
- Disease-group: developed for a group of similar conditions (e.g. cancer, neuromuscular diseases). They may be particularly valuable for rare diseases where it may not be feasible to validate a disease specific PROM given small patient numbers.
- Symptom-specific: focussing on a symptom, such as anxiety and depression as used in the GARDIAN registry.

HTA seeks to understand the consequences of using a new medicine vs using the standard of care, with two main approaches used in different countries: added benefit assessment and/or cost-effectiveness assessment. Both approaches use quality of life to understand the impacts of a treatment on patients, but the way that this is done in these approaches differs.

In added benefit assessments, PROMs are considered by some countries as an important additional endpoint to show what difference the treatment makes to patients in practical terms and may be measured as a continuous variable (change from baseline to end of treatment) or with the expectation of achieving a certain improvement (responder).

In cost-effectiveness evaluations, health effects are quantified using a composite measure of length of life adjusted for quality of life, called “quality-adjusted life years” (QALY). Within the QALY, quality of life is measured using a numerical score that is derived from PRO data. This score is referred to as a health state utility value (HSUV), which ranges on a scale of 0.0, representing dead, to 1.0, representing full health. In most cases, the EQ5D is the preferred PROM for adults as it is a generic PROM comparable across diseases and HSUVs can be easily calculated by averaging the sum of combination of individual responses, weighed based on public preferences (that have been collected by many health systems). If EQ-5D is not used, HSUVs can also be derived by statistically mapping other PROM responses to EQ-5D preference measures. Alternatively, vignettes can be used to describe health states of a condition and support direct elicitation of preferences.

As the previous speakers have outlined, measuring the burden of disease and finding appropriate clinical outcomes to measure the effects of an intervention is more difficult in rare diseases, and so quality of life data are more important. However, appropriate PROMs are difficult to identify, develop and validate in such small, heterogeneous, populations. Additional complications arise in paediatric populations. Furthermore, it is difficult to recruit sufficient patients over a long enough timeframe to enable an analysis that is adequately powered to demonstrate treatment effects. For this reason, analyses of quality of life data often show non-significant benefits and this evidence is criticized by HTA bodies.

Hence the IMPACT HTA research was constructed to develop guidance on how quality of life evidence could contribute better to HTAs, either in added benefit assessments or in cost effectiveness.

8 <https://euroqol.org/>

9 <https://nexusipe.org/advancing/assessment-evaluation/short-form-sf-36-health-survey>

10 Powell PA et al. Development of a New Quality of Life Measure for Duchenne Muscular Dystrophy Using Mixed Methods: The DMD-QoL. *Neurology*. 2021; 96:e2438-e2450.
doi: 10.1212/WNL.00000000000011896.

It involved several strands of multi-methods research with systematic literature reviews and expert consultations about how evidence from PROMs was used in HTA of rare disease medicines, mapping processes for HSUVs, and how different HTA bodies valued and critiqued PRO evidence. This led to five recommendations¹¹, as presented in Table 1, that aim to support consideration of quality of life impacts in HTAs that would be fair for rare disease treatments.

Table 1. Recommendations for Improving Interpretation of Evidence Relating to QoL in HTA of RDTs¹⁰

1	Given the complexity, severity and lack of clinical knowledge associated with many rare diseases, it is essential to evaluate the impacts of the condition, and the impacts of treatments, on the quality of life of patients and carers
2	When critically assessing evidence, challenges related to development and administration of patient reported outcome measures in rare diseases should be taken into account
3	During the appraisal, interpretation of evidence from patient reported outcome measures and health state utility values should recognize that lack of significant effect does not necessarily imply lack of benefit on quality of life
4	Other forms of evidence, including non-conventional approaches to quality of life and patient and clinical input and evidence, should be considered to enable a fuller appreciation of the impact of a rare disease treatment on quality of life
5	It is important to consider quality of life impacts on the family and carers to better capture the added benefit of a medicine

1. *Given the complexity, severity and lack of clinical knowledge associated with many rare diseases, it is essential to evaluate the impacts of the condition, and the impacts of treatments, on the quality of life of patients and carers*

HTAs evaluate comparative/relative effectiveness and an essential first step of this should be to understand the impacts of a condition on a patient under the current care pathway. However, from the documentation of appraisal processes in Europe, Canada and Australia¹² it is unclear whether burden of disease is specifically considered in many jurisdictions. This is essential for complex rare diseases that can have many clinical manifestations, can affect childhood development, may require support in a variety of ways and may progress over many years. As highlighted in an analysis of 24 non-cancer orphan drugs appraised by NICE, the quality of life of patients (and the majority of carers) were affected by the condition they are living with, and most treatments intended to improve their QoL but not all their length of life¹³.

To help understand the impact of treatment on quality of life, PRO data need to be collected routinely and considered in all appraisals, particularly for chronic rare diseases where improving length of life may not be the goal. The impacts of living with a condition can be shown from registries such as GARDIAN that collect a range of PROMs over many years and through other bespoke questionnaires.

¹¹ Nicod, E et al. Improving Interpretation of Evidence Relating to Quality of Life in Health Technology Assessments of Rare Disease Treatments. *The Patient - Patient-Centered Outcomes Research* 2022; 16: 7-17.

¹² <https://www.impact-hta.eu/country-vignettes>

¹³ Nicod, E et al. Consideration of quality of life in the health technology assessments of rare disease treatments. *Eur J Health Econ* 2022; 23: 645–669. <https://doi.org/10.1007/s10198-021-01387-w>

If such information is not available for an HTA submission, the input of people living with the condition and clinical experts as part of the deliberative HTA process is important. Allowing for patient and clinical expert input at the beginning of the appraisal meetings can be particularly useful in providing useful insights for the HTA decision context. To aid this, patient groups can support routine, long-term PRO data collection, as outlined by the IGA and several other presentations by patient groups in this HTAi conference.

Health Technology Developers need to produce the best possible evidence about quality of life impacts, by ensuring development and selection of PROs capture what matter most to patients, and meet the objectives of multiple stakeholders and then using them in research that pre-specifies what is a meaningful treatment benefit and powering a study appropriately.

2. When critically assessing evidence, challenges related to development and administration of patient reported outcome measures in rare diseases should be taken into account

Selecting an existing PROM or developing a new one is more difficult when disease progression and natural history are not well understood, when diseases have many clinical manifestations, and it is unclear what outcomes matter to patients. When developing a PROM there may be difficulties in identifying patients to be involved in the initial qualitative research, in demonstrating concept validity. Multi-site and international studies are often needed to generate a larger sample size, but confirming the psychometric equivalence of the PROM between countries may prove challenging.

The collection of sufficient PRO data from rare disease patients can be challenging because the patients may be children or have cognitive, motor or speech deficits that affect ability to obtain information. Proxy-reporting from carers and families is common, but may not truly reflect patients' issues. Even for more prevalent conditions there is often a large amount of missing data in PRO evidence that may not be at random (either missing because improved, or worsened). This is a challenge for a smaller, rare disease population and so extra care should be taken to support regular completion of PROMs. Well-informed site staff and patients and active monitoring of PRO compliance can help limit missing data.

Given these challenges there is a risk that PRO evidence does not show significant or clinically meaningful results and there are a range of issues that further affect the analysis. PROMs may be subject to ceiling and floor effects. The early onset and severe nature of many rare diseases can produce a degree of response shift whereby people living with the disease adapt to their condition and consider it as "their normal". In paediatric conditions, the measurement of the burden of disease and benefits of treatment is complicated by the child's natural development. The child's development (even if delayed) can suggest gains in QoL as they develop physically, intellectually and socially, and this makes comparison against baseline misleading. As a result, generic measures or other condition-specific PROMs may be used, but there are concerns about content and face validity. Furthermore, PROMs may not sufficiently sensitive due to insufficient power or length of study.

These aspects should all be explained when PRO evidence is submitted and taken account of in assessment.

3. During the appraisal, interpretation of evidence from patient reported outcome measures and health state utility values should recognize that lack of significant effect does not necessarily imply lack of benefit on quality of life

HTA bodies have different evidentiary requirements for PRO evidence. Some bodies (such as the G-BA in Germany) require specification of a pre-defined minimally important difference, which may not be appropriate for heterogeneous rare diseases. Some HTAs impose analytical requirements (such as requiring the endpoint to be validated, or a co-primary or endpoint hierarchically tested – HAS in France), which may lead to diverging appraisals of the same evidence. For example in the appraisal of inotersen for amyloidosis, there were different interpretations of the statistically significant Norfolk Quality of Life-Diabetic Neuropathy data. The G-BA classified the effect as inconclusive due to lack of clinical relevance, whilst HAS concluded modest improvement in this co-primary endpoint.

In a review of HTAs of rare disease medicines, PROMs tended to be used to study (groups of) symptoms, rather than quality of life as a whole, which is the focus of HTA bodies. The reason for this is perhaps due to the previously stated challenges, but the rationale for selection of PROMs and how they should be interpreted needs better explanation from Health Technology Developers. Furthermore, Health Technology Developers need to consider the needs of HTA bodies (not just regulators) when designing their trials and all need to work together to align requirements.

For cost effectiveness analyses in rare diseases, an extra layer of complexity is added. Not only is the PRO data translated into a numerical HSUV, but the data must fit the specific health states included in the model. If HSUVs are derived from EQ-5D, the translation into a numerical value is straightforward but there are concerns about the sensitivity of the generic measure and whether it appropriately captures the impacts of disease and treatments. If mapping of PRO data to HSUVs is required, large and representative datasets are needed and prone to measurement error and high levels of uncertainty, and in some cases may not be possible at all. It is also more difficult to produce QoL data that also fits the health states included in the model, in which cases other data is being used which then also comes with some additional uncertainty.

There have been cases where HSUVs gave completely different conclusions on QoL impacts than have been observed, depending on the definition of health states and approach used to derive the value. As such, decision-makers need to consider the validity of the way in which the QALY has been constructed and the viability of the economic model for decision making purposes taking account of clinical and patient expert opinion.

4. Other forms of evidence, including non-conventional approaches to quality of life and patient and clinical input and evidence, should be considered to enable a fuller appreciation of the impact of a rare disease treatment on quality of life

When conventional approaches for generating quality of life evidence are not feasible, published literature from a similar disease could be used. This may be acceptable if clinical and patient experts confirm that the proxy disease is similar in nature and patient population (e.g. age and presentation of disease). In economic modelling, particular consideration must be given to the similarity of health states.

For HSUVs, vignettes are used more frequently in HTAs of rare disease treatments. Vignettes of each health state need to be carefully developed through robust qualitative research with people living with the disease and clinical experts to assess content and face validity. Otherwise HTA bodies may have concerns about bias (assigning lower utilities to later disease states) and so careful design and reporting of these studies is required.

Health Technology Developers should be encouraged to undertake qualitative research to develop robust patient-based evidence for HTA purposes that can be used to explain the burden of disease and for example use exit interviews in clinical trials to capture treatment impacts that are not measured in standard endpoints. Ideally the general research on burden of disease should be co-created with patients in a pre-competitive manner involving a range of developers and academics that understand HTA, such as in Project Hercules¹⁴. In any case, it should be published in a peer-reviewed journal to increase community understanding of the disease.

The extent to which patient and clinician input are influential in appraisal depends on the HTA programme and its processes for stakeholder involvement. This includes mechanisms for submission of input from people living with the disease and clinical experts, the form of participation in the appraisal meeting and trust in the input, given perceived conflicts of interest. Patient and clinical groups may draw on surveys, patient stories, interviews or focus groups to develop their submission. Their input can help confirm or refute inconclusive results from quality of life evidence in the manufacturer's submission and inform discussion about whether the underlying states in an economic model are suitable for decision-making. As a result, HTA reports should document the patient and clinical inputs and highlight those aspects that impacted decision making relating to burden of disease, preferred outcome measures, the adequacy of PRO instruments, and HSUV derivations. For example, whether the chosen instrument captures their experiences and has face validity (e.g. losing vision should not equate to a state worse than death).

5. It is important to consider quality of life impacts on the family and carers to better capture the added benefit of a medicine

Rare diseases that affect children, and those that have rapid progression later in life are extremely burdensome for carers, in terms of the time needed to provide personal care for the patient, administer therapies and travel to clinic visits. This may require reduction in working hours, have a variety of impacts on physical and mental health, and added financial burdens may be placed on a family when there is reduced income. Additionally, there is often uncertainty about when and how patients may deteriorate, which increases stress levels for patients and family members. Parents may also experience grief and desperation when caring for young children or infants who may be severely debilitated, or for whom there may be little hope of survival. The genetic nature of these conditions also means that siblings may be affected and may be at different stages of diseases creating different demands on parents.

So HTA of rare disease medicines should include impacts on carers (and their families). However, there are no standard methods for doing this and so if carer impacts are considered there may be inconsistency of approach (e.g. NICE has used different decrements in utility to take account of carer burden depending on age and severity of condition and made a judgement about the level and number of carers needed, but not in a consistent approach). The need for development of research in this area has been recognised and is a high priority for rare diseases.

¹⁴ <https://hercules.duchenneuk.org/>

Conclusion

These recommendations highlight the challenges using PROMs in rare diseases to determine the impact of the condition on people living with the disease and the impacts of a treatment on a patient and a carer's quality of life. These lead to key actions for all stakeholders:

- Health Technology Developers, academics and patient groups: co-create and publish research that describes the burden of disease and its progression over time in an open and transparent manner; collaborate early in drug development and work towards collecting data that meet the needs of all key stakeholder, including HTA
- Health Technology Developer: in an HTA submission explain the rationale for use of PROMs in clinical studies and limitations in analysis; presents other patient-based evidence to show treatment impacts on quality of life.
- HTA bodies: contribute to joint scientific advice processes to agree what quality of life evidence can be collected; assess what is feasible with PRO evidence in the specific rare disease being assessed; consider different analytical approaches to measuring quality of life impacts; involve patient and clinical experts in the deliberative process to explain burden of disease and help interpret impacts of the treatment on quality of life that may not be captured in the main HTA analysis; critically assess whether the economic model based on a QALY is sufficiently robust for decision-making; support methodological developments – e.g. in assessment of carer quality of life.

4. A Clinician's Perspective on a New Paradigm for Evidence Generation of Patient Benefit

Alfonso Iorio PhD - McMaster University, Canada

Professor Iorio, built on Dr Nicod's recommendations about how a wide range of evidence pieces from different stakeholders could be used to demonstrate the impact on quality of life of a rare disease and its treatment. He provided a clinician's view on the limitations and opportunities of the current approach to HTA for rare disease treatments, the potential for different stakeholders to work together differently and in partnership, and the role of digital technologies to support patient co-creation of evidence.

Developing the HTA paradigm for rare diseases

Clinicians, patients and health technology assessors all have concerns about the sustainability of health systems, as most systems try to provide universal health coverage with limited budgets, shortages of staff, long waiting lists and increasing provision of innovative technologies at a high cost. It is intuitively acknowledged that the needs of an individual patient living with a rare disease must be balanced against the needs of all other users of the healthcare system, to ensure that the system is fair for all, that health expenditure cannot continuously increase. As a result, value propositions are urgently needed to allow opportunity choices for health system expenditures.

The current paradigm for assessing the value of treatments for rare diseases does not work. It seems to be like a square peg (of evidence for rare disease treatments) is smashed into a round hole (of HTA), and in doing so the whole health system infrastructure is becoming fractured and unsustainable. To resolve this, there needs to be a paradigm shift in the appraisal of clinical effectiveness to ensure optimal care for all.

In terms of the task of fitting the square peg into the round hole, as often happens, changing the square peg into a round one (changing the evidence generated in a rare disease), or making the round hole squared (changing the HTA rules around evidence generation) are not feasible, let alone optimal solutions. The solution is likely to be to find a polygonal shape, for both the peg and the hole, that better fits the ultimate goal of providing cost-effective interventions to all those in need. One effective way to think of this polygon is to envision it as a table around which all the involved partners undertake a conversation to ensure everyone's goal is achieved.

Stakeholder partnerships and responsibilities

Stakeholder involvement has been repeatedly shown to be unsuccessful when sought under a “paternalistic” assumption where only one of the stakeholders (be it industry, regulators, or researchers) drives the agenda seeking support from other stakeholders (usually patients) to achieve their own goal. Successful stakeholder involvement requires the early formalization of a partnership, where stakeholders come around the planning table with the opportunity to contribute to shaping a common goal.

Ms Collin-Histed's presentation was a compelling, complete demonstration of how this can be accomplished in real life. The definition and adoption of the CoreHEM outcome set¹⁵ for gene therapy in haemophilia is another example of successful partnership covering the entire arch of development of an intervention from bench to coverage decisions.

With this context, the following suggests a few aspects of what different stakeholders may bring, or take away, from sitting around a common table, with the caveat that this is written from the view of a clinician involved in a range multi-stakeholder collaboration.

Clinicians

Clinicians do not usually think about the HTA question of the value of a new treatment within the perspective of a health system; where “value” has to be appraised and quantified in a way conducive to translation to a cost the system is willing to pay. Indeed, clinicians and researchers usually focus on how they can help bring better treatment options to patients. They understand the need to measure impact, both of the untreated disease and new treatments, but usually focus on outcome measures related to the pathophysiology they are trying to modulate (e.g. improvement of laboratory values, or sophisticated measures of body functioning), or at best on hard clinical outcomes (e.g. survival, disease progression). So, they may not drive for the inclusion of measures of QoL or burden of disease measures, but they often allow and support their inclusion in the study design and routine clinical care (including recording in registries) when working in partnership with patients.

¹⁵ Iorio, A., Skinner, M. W., Clearfield, E. et al. (2018). Core outcome set for gene therapy in haemophilia: Results of the coreHEM multistakeholder project. *Haemophilia*, 24(4), e167–e172. <https://doi.org/10.1111/hae.13504>

Patients

The evolution in our collective way of thinking about patient involvement over the last few years has changed thanks. In North America, this has been mainly due to impactful initiatives like PCORI in the US or SPORs in Canada. Partnership with other stakeholders, clarifies to patients, the importance, and implications, of their involvement in evidence generation. It can also highlight that pragmatic research on lifelong diseases will last for many years and requires enrollment of “most” patients as part of an endless cycle of drug development. This will have a “human cost” for many patients to enable (very) future patients to enjoy the value of the treatment that they are allowing to be tested on themselves. The implications are huge, as patients will not remain on the receiving end of goods with large societal costs. The traditional model where Health Technology Developers bear the cost of developing a new treatment to then recover it, producing revenues, may not work for rare diseases. Research performed during a progressive, managed access to care would (when accepted by all partners) offers an opportunity to (partly) offset the cost of research and development by direct involvement of patient and care givers in data generation.

Health Technology Developers

Further to the considerations about patient involvement, partnership may imply the need to shift the typical risk sharing attitude to including payment modalities, repayments when the treatment does not work or stop working, etc. Furthermore, it is important to acknowledge as early as possible in an open conversation with all the stakeholders that, irrespectively of the size of the affected population, what will be paid for by health systems is the “health value”, which may be insufficient to re-pay the research and development. Perhaps health care systems would be ready to entirely re-pay the research and development costs for a life-saving cure for a deadly disease with no available options, but in many other cases alternative business models to support development and delivery may be needed.

Furthermore, measures assessing value should be incorporated as early as possible in the development of the intervention, to avoid wasting resources on developing interventions that may never demonstrate enough value for HTA/Payers.

Health Technology Assessment

All agree that multi-stakeholder scientific advice (including regulators and HTA bodies) is the most valuable tool to support dialogue about optimal pivotal study design and this is particularly true for rare diseases. However, there are concerns about the limited capacity of such processes to cover all treatments in development and to enable iterative dialogues over the life cycle of development, which is particularly important in rare diseases where knowledge evolves.

There is growing understanding that there will be higher uncertainty in the demonstration of effects from rare disease treatments compared to more common conditions. Even when state of the art RCTs could be performed, they would be small and extremely long, to ensure measurement of burden of disease and impact on QoL, and so it is unlikely that they will be undertaken. Evidence generated via registries, mostly providing before-after comparisons, may be generated and collected minimizing biases, but will always have the potential for high uncertainty. This uncertainty needs to be factored in and may translate into a “discount” of what is expected. This is not a “discount” (increasing) on the QALY threshold, but rather a discount of the level of certainty around the effect of the interventions.

There is nothing magical about a p-value of 0.05 and one could easily envision accepting different p-values, to be set and refined over time, disease by disease. Then discounts on the certainty threshold would translate into discounts on the price paid for the intervention and managed access to it, at least until more evidence could be generated.

Dealing with uncertainty requires navigation tools. The impact of unavoidable uncertainty cannot be loaded only onto the patients, clinicians and Health Technology Developers: HTA bodies, if willing to enter a partnership framework, will need to get involved and issue specific guidance on what evidence they will accept for a specific disease, within a specific development plan. As regulators often do, HTA bodies will need to sit at the planning table, perform appropriate modelling of what is possible to measure and achieve within the boundaries of that specific disease and research plan, and declare upfront what they would need and be able to accept for decision making.

A societal perspective may need to be taken on the entire process of developing and vetting interventions for rare diseases, where the “cost” for the application and review of interventions for rare disease charged to the Health Technology Developer is lowered (with respect to applications for common diseases) depending on stage of research, the size of the affected population, and the amount of evidence to be reviewed. In a societal perspective, it is not impossible to imagine a system charging (part of) the cost incurred for appraising interventions for rare disease on applications for interventions for larger markets. Finally, in a partnership framework, HTA reviews may become “adaptive”, incorporating new evidence as it comes and with evidence generation plans embedded in managed entry agreements.

Payers

For payers, once it is agreed that “health value” is what is monetized, it would be important to work in partnership with all the others to devise innovative ways of assessing how and which value is explored, including agreed approaches to making health values in different disease areas relative to each other, thus allowing harmonic, equitable, and sustainable growth of the entire envelope.

Overview of stakeholder roles

A partnership approach would prompt all stakeholders to think and act differently, co-designing solutions that will lead to generation of better evidence for rare diseases. This should start with, and maybe be “chaired by”, those who live 24/7 with a condition and know its impacts, the people living with the disease. They should be encouraged to, and supported in, generating their own data that reflects how they perceive the progression of their disease and how they feel about their quality of life.

Such an approach enables knowledge sharing, supports mutual learning, and builds empowered communities that can deliver powerful evidence-based advocacy to a range of decision makers, as so clearly outlined by IGA in Ms Collin-Histed’s presentation. Such approaches usually require large commitments in terms of time, money and psychological burden and so should not be undertaken without careful planning. These approaches require careful protection from the risk of being “hijacked” by individuals or stakeholders with strong pre-conceptions about what potential solutions are, thereby losing the openness to evaluate impacts and value in an unbiased manner.

Digital technologies

A final set of considerations relates to the role of technology in patient-led co-creation approaches to evidence generation. The massive diffusion of wearable devices has changed the prospects for capturing patient reported outcomes ^{16, 17}.

Wearable devices can be complicated, research-oriented devices with sensors collecting sophisticated information on human metabolisms and pathophysiology. However, for burden of disease, and quality of life, the most effective and attractive wearable device is a smart phone. Almost every individual on earth, irrespective of demographic or health condition has, and often wears on his body, a smart phone. More and more have a wrist device capable of collecting blood pressure, pulse, temperature, fine movements, sleep or awake status, oxygen saturation. Furthermore, major manufacturers of smart phones have created and released infrastructures like Apple or Google Health, allowing sharing of raw or computed data collected by the many sensors, including the accelerometer, for research use.

It is helpful to consider a few examples of apps that have been created and validated to measure patient important outcomes.

The Apple iWatch can identify bouts of atrial fibrillation, alerting the individual, his caregiver, or his doctor¹⁸. The application was criticized for not being very sensitive, but was found to be very specific, meaning that when compared with a week-long recording of an EKG (Holter recording), it missed about 31% of cases. However, 81% of those identified were actual cases. This technology is clearly not ready to substitute the Holter monitor, but it can still save many strokes by correctly identifying more than half (approximately 54%) of the cases of atrial fibrillation that would otherwise go undiagnosed and therefore untreated.

The Apple iPhone can consistently measure activity as precisely as the 6-minute walk test (6MWT)¹⁹, and can do that many times a day for as long as one wants, at no cost. Using such apps to measure impacts on QoL in the real-world setting would be extremely valuable. One consideration in this case is that the application was found to correlate well with the 6MWT. Had it not, one would seriously challenge the validity of the traditional 6MWT. This implies that we need to be ready to change our concept of gold standard measurement. If the iPhone app can distinguish different levels of activity in different individuals, and in the same individual over time, and a clinically performed artificial assessment cannot, which one should we think is more applicable to reflecting real life activity levels? An imprecise measure averaged over thousands of recordings during real life conditions, or an artificial measurement made once or twice in a clinical assessment setting. The answer will only be found by prospectively assessing which one correlates better with hard outcomes (e.g. survival, cardiovascular events, falls, etc). Furthermore, the app can quickly detect variations in the steadiness, pace and rhythm of walking of the individual, which have been shown to correlate with the risk of subsequent falls²⁰.

¹⁶ Lammers, A. E. et al. (2022). Reference Values for Wristband Accelerometry Data in Children Aged 6–11 Years of Age. *Frontiers in Pediatrics*, 10. <https://doi.org/10.3389/fped.2022.808372>

¹⁷ Hirsch, M. C. et al. (2020). Rare diseases 2030: How augmented AI will support diagnosis and treatment of rare diseases in the future. *Annals of the Rheumatic Diseases*. <https://doi.org/10.1136/annrheumdis-2020-217125>

¹⁸ Miranda-Arboleda, A. F., & Baranchuk, A. (2022). Atrial Fibrillation Detection by Smartwatch Devices in Patients With Underlying ECG Abnormalities: Still Not Smart Enough? *Can J Cardiology*; 38:1713–4.

¹⁹ Mak, J., et al. (2021). Reliability and repeatability of a smartphone-based 6-min walk test as a patient-centred outcome measure. *European Heart Journal - Digital Health*, 2(1), 77–87. <https://doi.org/10.1093/ehjdh/ztab018>

²⁰ Rubin, D. S., Ranjeva, S. L., Urbanek, J. K., et al. (2022). Smartphone-Based Gait Cadence to Identify Older Adults with Decreased Functional Capacity. *Digital Biomarkers*, 6(2), 61–70. <https://doi.org/10.1159/000525344>

One final example is about AI applications elaborating in real time information (including videos) collected by a system of sensors in the living space of an elderly patient²¹. The AI component can quickly and very precisely identify variations in the time from lying in bed to standing up, “normalcy” of trajectories in the space, longer than usual idle periods, and generate alarms for a supervisor. Similar systems have been successfully applied by Kaiser Permanente for monitoring ICU beds, or newborns, and are used by selected living facilities for elderly people. They could easily be adapted to support people living with rare diseases at home.

Privacy, consent, freedom are of course important considerations; however, patients and carers are enthusiastic about the opportunity, and welcome the technology for the benefits it provides. On the other hand, are not we all sharing our positions and movements with everyone else when we use Apple or Google Maps? Are we not buying on Amazon? Calling Ubers? We are all sharing data out of our wearable technology, and often much more than we realized, until we recognise that we are getting targeted by very tailored advertising from local or global dealers of articles we may be interested in.

New data governance models are needed whereby patients can share their own data with researchers, health technology assessors and policy makers in a secure way, and we need to explain the benefits of this secondary use of data. However, no one should doubt this will happen in the very near future. There is no ethical or privacy barrier able to stop the progress of technology, and particularly of AI. The review of times to 1 million users for web-based applications speaks for itself: it was 3.5 years for Netflix, 2.5 years for Airbnb, 2 years for twitter, 10 months for Facebook, 5 months for Spotify and 5 days for ChatGPT!

AI and digital technologies have finally opened up a new way of collecting plenty of reliable patient relevant outcome measures. Patients will generate the data, and it is now up to us, researchers and health technology assessors, to learn how to best leverage these data to support the needs for effective treatments of patients with rare diseases.

5. Discussion

There was wide-ranging discussion with the audience about the challenges of generating robust evidence from PROMs in rare diseases including:

Challenges in evidence generation:

- difficulties in collecting data from paediatric patients or patients with cognitive impairment
- subjectivity of response that may be affected by something that is not directly related to the rare disorder, such as time of day, mood, having an intercurrent illness or accident, etc
- the importance of measuring caregiver quality of life, but the limited number of instruments available
- the potential value of data from apps and AI that automatically capture changes in patient functioning, but the potential challenges of validation (even in larger populations) and acceptance as a new source of evidence by decision makers.

²¹ apps.apple.com/us/app/care-daily-ai/id1454869628

Paediatric populations

- increased uncertainty, linked to lack of validation of PROM
- potential biases with proxy measurement by carers
- how to give more power to e.g. KOLs to validate soft data.

Data ownership:

- Who owns the data collected – the patients or the clinicians or a 3rd party?
- Careful consideration needed about protecting data confidentiality given the risk of de-anonymisation in small populations.
- Who can make decisions on how data is used and when it is withheld (e.g. to safeguard a publication)?

Overall it was felt that it is more important to understand quality of life issues for patients with rare disorders and their carers, but that there are major measurement and analytical challenges that need further work. The input of expert patients and clinicians to help develop valid measures and interpret data is essential.