Challenges in Research and Health Technology Assessment of Rare Disease Technologies: Report of the ISPOR Rare Disease Special Interest Group

Sandra Nestler-Parr, PhD, MPhil, MSc1, Daria Korchagina, PhD, MSc2*, Mondher Toumi, MD, MSc, PhD3, Chris L. Pashos, PhD4, Christopher Blanchette, PhD, MBA5, Elizabeth Molsen, RN6, Thomas Morel, MSc, MA6,7, Steven Simoens, PhD, MSc, MA8, Zoltán Kaló, PhD8,9, Ruediger Gattermann, MA, MBA10, William Redekop, PhD11

1Rare Access Ltd, Leamington Spa, England, UK; 2Mental Health and Public Health Unit (Inserm U669), University of Paris-Sud, Paris, France; 3Public Health and Chronic Disease Laboratory, Aix-Marseille University, Marseille, France; 4Global Outcomes and Epidemiology Research, Takeda Pharmaceuticals International, Inc., Cambridge, MA, USA; 5College of Health and Human Services, University of North Carolina at Charlotte, Charlotte, NC, USA; 6Scientific & Health Policy Initiatives, ISPOR, Lawrenceville, NJ, USA; 7KU Leuven Department of Pharmaceutical and Pharmacological Sciences, Leuven, Belgium; 8Department of Health Policy and Health Economics, Eötvös Loránd University (ELTE), Budapest, Hungary; 9Syreon Research Institute, Budapest, Hungary; 10Healthcare Policy and External Affairs Europe, CSL Behring, Biotherapies for Life, Marburg, Germany; 11Health Technology Assessment, Erasmus University, Rotterdam, The Netherlands

ABSTRACT

Background: Successful development of new treatments for rare diseases (RDS) and their sustainable patient access require overcoming a series of challenges related to research and health technology assessment (HTA). These impediments, which may be unique to RDS or also apply to common diseases but are particularly pertinent in RDS, are diverse and interrelated. Objective: To develop for the first time a catalog of primary impediments to RD research and HTA, and to describe the cause and effect of individual challenges. Methods: Challenges were identified by an international 22-person expert working group and qualitative outreach to colleagues with relevant expertise. A broad range of stakeholder perspectives is represented. Draft results were presented at annual European and North American ISPOR congresses, and written comments were received by the 385-strong ISPOR Rare Disease Review Group from two rounds of review. Findings were refined and confirmed via targeted literature search. Results: Research-related challenges linked to the low prevalence of RDS were categorized into those pertaining to disease recognition and diagnosis, evaluation of treatment effect, and patient recruitment for clinical research. HTA-related challenges were classified into issues relating to the lack of a tailored HTA method for RD treatments and uncertainty for HTA agencies and health care payers. Conclusions: Identifying and highlighting diverse, but interrelated, key challenges in RD research and HTA is an essential first step toward developing implementable and sustainable solutions. A collaborative multisectoral effort is required to enable faster and less costly development of safe, efficacious, and appropriate new RD therapies that offer value for money.

Keywords: cost-effectiveness, health policy, health technology assessment, orphan designation, rare diseases.

Background to the Rare Disease Working Group

In 2013, two working groups were established under the auspices of the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Rare Disease Special Interest Group. The first working group undertook a review of rare disease (RD) terms and definitions, motivated by recognition of the lack of a universal definition of rare diseases or health technologies for their treatment and the existing diversity of definitions applied to rare diseases. The output of this research was the article “Rare Disease Terminology and Definitions—A Systematic Global Review: Report of the ISPOR Rare Disease Special Interest Group,” published in 2015 in Value in Health [1]. The second working group, Challenges in Research and Health Technology Assessment of Rare Disease Technologies, identified and...
reviewed challenges faced by those engaged in research and health technology assessment (HTA) in RDs and their treatments. The goal of the working group was to evaluate these challenges and disseminate the findings via publication and presentations. An outline was initially developed by working group members representing different stakeholder perspectives from Europe and the United States, and sequential drafts were reviewed and modified during monthly teleconferences among the coauthors. Further feedback was obtained during work-to-date presentations given at annual ISPOR European congresses in Dublin, Amsterdam, Milan, and Glasgow, and annual ISPOR international meetings in Montreal and Boston. The final version of this article represents the outcome of these discussions, conference feedback, and written comments received by members of the ISPOR Rare Disease Special Interest Group from two rounds of review.

Introduction

Approximately 60 million people in the United States and European Union are affected by an RD [2]. Although RD research and clinical development of technologies for the treatment of RDs are rapidly expanding areas, there is still no universally accepted terminology or definition as to what constitutes an RD. Typically, RDs are characterized by low frequency, where frequency is expressed in terms of prevalence or incidence within a specific country or geographical region. A global review of RD terminology found that 58% of definitions included a prevalence threshold with an average global threshold of 40 cases/100,000 people [1].

According to the European Organisation for Rare Diseases (EURODIS), an alliance of more than 700 RD patient organizations in 65 countries, more than 6000 distinct RDs exist, of which approximately 80% are of genetic origin [3]. On average, five new RDs are described every week in the medical literature [3]. RDs represent a broad assortment of disorders and constellations of clinical signs and symptoms but the vast majority of RDs affect children and are chronic and life threatening [4]. No cure exists for the substantial majority of RDs, and only a few RD treatments with proven efficacy are currently available [5].

Consequently, countries throughout the world have recognized the need to enact laws and regulations to provide incentives for the development of new and innovative technologies for the treatment of RDs [6–10]. Advancements in molecular genetics, understanding of disease pathogenesis, and medical technology have led to enhanced identification of RDs and pathways for improving RD diagnosis, prognosis, and treatment, as well as more accurate subclassification of common diseases into collections of RDs with distinct phenotypes [11–13].

However, the development of new RD therapies faces significant obstacles with respect to research and HTA. These challenges had not previously been evaluated comprehensively and, consequently, this ISPOR Rare Disease working group developed a multistakeholder catalog of the principal difficulties faced in RD research, during evidence generation for HTA, and HTA of RD treatments. Although identification of the obstacles is an important first step toward providing efficient and practical solutions, the working group anticipates the generation of another detailed report providing recommendations to address the challenges identified here. Challenges related to pricing of health technologies for RDs, their adoption, and patient access were not within the remit of this project.

Methods

Impediments to RD research and HTA were identified by a working group comprising 22 members with relevant expertise, as well as through qualitative outreach to colleagues specializing in RDs in contract research, the life sciences industry, and academia. The preliminary list of challenges underwent three rounds of review by the working group for comprehensiveness, refinement, and merger of duplicates. The challenges identified were analyzed for interrelationships and classified into categories. Findings were underpinned by a targeted literature search. Written comments were received by the 385-person ISPOR Rare Disease Review Group from two rounds of review and further verified when draft reports were presented at ISPOR annual international congresses in North America and Europe. A broad range of stakeholder perspectives from researchers, clinicians, industry, regulatory and HTA agencies, patients, payers, and market access specialists are represented in this report.

Results

The following sections describe the cause and effect of individual challenges and their relevance in RDs. Results are grouped into challenges relating to RD research and HTA of RD treatments, respectively, and subcategorized. It should be noted that, although the identified impediments are diverse, they are interrelated. Furthermore, some of the identified issues are unique to RDs, whereas others also apply to common diseases but are especially relevant or burdensome in RDs.

Challenges in Research

Research-related challenges linked to the low prevalence of RDs were grouped into three categories, as illustrated in Figure 1.

Disease recognition and diagnosis

Several interrelated challenges pertain to the recognition and diagnosis of RDs, all of which impinge on the quality of epidemiologic and clinical studies and complicate the characterization of unmet patient needs, potential efficacy, safety, effectiveness, and value of treatments for RDs.

Lack of familiarity with RDs. Insufficient awareness and knowledge of RDs can increase the likelihood of misdiagnosis and delayed accurate diagnosis [14–21]. Patients unfamiliar with pertinent signs and symptoms may not seek medical advice when appropriate. Similarly, clinicians may fail to recognize the disease [14,15] or may incorrectly attribute symptoms to common diseases with which they are more familiar. This is reflected in the average delay of 7.6 years in the United States and 5.6 years in the United Kingdom before a patient with an RD receives the correct diagnosis [22]. In a survey-based outcome study of symptomatic patients with a1-antitrypsin deficiency, an underrecognized rare genetic condition that increases the risk of lung emphysema and liver disease, the average diagnostic delay was 8.3 ± 6.9 years after onset of symptoms [23].

Disease heterogeneity. Incomplete understanding of a disease and its etiology may severely limit the comparability of findings from epidemiologic and clinical studies. Heterogeneity in pathogenesis, symptom presentation, natural history, disease severity, and progression can greatly impede efforts to characterize an RD in clinical research and to identify it in routine clinical practice, often resulting in misdiagnosis and an underestimation of true disease frequency.

The heterogeneous clinical presentation of many RDs hampers identification of affected patients, as seen across the broad spectrum of phenotypes in Gaucher’s disease, ranging from lethal disease in neonates to asymptomatic older adults [24]. It is not uncommon for patients with RDs, such as Behçet’s disease or late-onset Pompe disease, to exhibit a long initial asymptomatic phase during which their condition may not be identified [25,26]. In many RDs, no genotype–phenotype correlations have yet been
Different causations. The disease, initially termed because of similar clinical presentations were later found to have treatments for several syndromes within this disease family [30]. Investment over the past decade into development of innovative genetic origin of different MPS types have motivated substantial common disease mechanism [29,30]. The similarities in the lysosome enabled an accurate characterization of the underlying, to date were described and named before the discovery of the causality, as occurred in several genetic metabolic disorders unrelated have subsequently been found to have common between regions and countries, ranging from clustering in some geographic areas to wide dispersal in others [25,37,38]. A

In contrast, other diseases that initially seemed related because of similar clinical presentations were later found to have different causations. The disease, initially termed “weisses Blut” (German for white blood) by the German pathologist Rudolph Virchow in the 19th century who found abnormally high levels of white blood cells in his patients, is now known as “Leukämie” (Greek “leukos” for white and “aima” for blood) [31]. Today, leukemia is no longer seen as a single disease but recognized as a family of distinct diseases of different pathogenesis, which require markedly different therapeutic strategies [32].

Difficulties in establishing specific and sensitive diagnostic criteria. Diagnosis of an RD may be straightforward in patients with pathognomonic clinical features. However, owing to the heterogeneity of many RDs and the difficulty with correct interpretation of complex investigational testing algorithms, diagnostic certainty can be problematic. Establishment of relevant and specific diagnostic criteria, including diagnostic tests, may be hindered by the lack of (1) a sufficiently large patient cohort from which criteria can be reliably characterized, (2) consensus on diagnostic criteria due to heterogeneous disease presentation, and (3) diagnostic criteria that are applicable to all cases across the range of disease heterogeneity. For example, establishing the diagnosis of cystic fibrosis is straightforward in most cases, but it is more complex in a small proportion of patients. The “sweat test” to assess the chloride content of sweat is a standard screening test in infants but is often unreliable in patients with a less severe clinical phenotype who may present later in life [33]. Diagnosis through genotyping may also be problematic in milder cases, such as in the absence of a homozygous ΔF508 gene abnormality and when uncharacterized abnormal genes are present [33].

Misdiagnosis. The foregoing problems commonly result in misdiagnoses that can lead to inappropriate treatment. Such treatment may not only produce side effects, but can even mask symptoms of the underlying condition, thereby further delaying the correct diagnosis and the initiation of appropriate treatment. In α-1 antitrypsin deficiency, for example, patients are frequently misdiagnosed as suffering from asthma and inappropriately treated with inhaled asthma therapies. It may take several years of inadequate response to these treatments and multiple consultations with different specialists before the correct diagnosis is made [23,34]. In Wilson’s disease, a rare genetic disease of copper metabolism, the heterogeneity of initial symptoms and the broad age range at onset prompt presentation to a diverse range of specialists, which is a contributing factor in the diagnostic delay and frequent initial misdiagnosis of the condition [35]. The most common cause of death among patients affected with Wilson’s disease, which is always fatal if left untreated, is delayed diagnosis [36].

Geographic variation. The prevalence of an RD may vary between regions and countries, ranging from clustering in some geographic areas to wide dispersal in others [25,37,38]. A
prevalence study of Behçet’s disease, a chronic, multisystem vasculitis also referred to as “Silk Road disease,” revealed a significant geographic variation globally. The highest prevalence was reported in Turkey, at 421 cases/100,000 people, followed by other countries along the ancient trading route between the Far East and the Mediterranean Sea, and the lowest in the United Kingdom, at 0.6 cases/100,000 people [25]. Other examples of RDs that exhibit significant geographic variation in prevalence include Finnish disease heritage, a group of rare hereditary diseases that are overrepresented in Finland [39], and α-1 antitrypsin deficiency, in which the prevalence of the Z genetic variant that is associated with severe deficiency is considered to reflect the migration patterns of Viking descendants [40].

When the true geographic variation in disease frequency is unknown, prevalence data from epidemiologic studies performed in a particular region may be wrongly extrapolated to a wider geography in the event of prevalence heterogeneity. The existence of single-site or single-region disease cohorts may lead researchers to suspect the existence of regional variation in disease frequency [41]. Conversely, geographic differences in disease definition and diagnostic criteria may result in over-estimates of regional prevalence variation. In Behçet’s disease, for example, a higher percentage of cases in low-prevalence countries has been seen in patients whose ancestry is traceable to high-prevalence areas [25]. However, the disease frequency in these isolated high-prevalence pockets is not generalizable. Although robust epidemiologic studies might establish the degree of phenotypic variation in disease frequency and thereby help to ascertain disease etiology, they may not provide sufficient data to allow the distinction of genetic from environmental factors, especially when conducted in populations that rarely migrate.

Evaluation of treatment effect
The reliable assessment of the effectiveness of medical interventions for the treatment of conditions across the entire spectrum of disease severity and prevalence presents many challenges. However, some of these challenges are of particular relevance in RDs as a direct consequence of their rarity and variable clinical phenotype, as detailed previously.

Heterogeneity of prognosis and treatment effect. The combination of disease rarity and heterogeneity reduces the certainty with which an RD technology’s efficacy, safety, and effectiveness can be determined, as treatment effect may vary significantly in a single RD [34-36]. This variability also hampers interventional studies as it requires increased sample sizes for suitably powered clinical trials [42,43].

Genetic and biological markers are potentially useful tools to identify patient subgroups that are most likely to respond to a particular treatment [44], to develop targeted treatment approaches, and to predict treatment responses and prognosis [45]. For example, several technologies were developed for cystic fibrosis that target specific mutations in the cystic fibrosis transmembrane conductance regulator (CFTR) gene [46-48]. However, the utilization of biological and clinical markers for patient recruitment into clinical trials will have conflicting effects on heterogeneity and rarity; although subgroups may be more homogeneous, they will de facto be smaller in size. This may substantially prolong patient recruitment in studies of treatment efficacy and safety.

Selection bias. Clinical trials may be biased toward the inclusion of patients with more severe disease and more severe symptoms, whereas patients with a milder disease course may either not be identified as having the RD in question or may not meet the inclusion criteria. Selection bias may also impede generalizability of the trial results to a broader patient population in other ways. For example, clinical trials of novel treatments for multiple myeloma often exclude elderly patients even though the incidence of myeloma is higher in the elderly [49]. Consequently, treatment efficacy demonstrated in controlled clinical trials may differ from the effectiveness observed in routine clinical practice (although this may not always be the case [50]).

Clinical trials tend to be conducted in centers with expertise in specific RDs, which may lead to selection bias toward those patients more able to withstand travel to the study center [51–53]. Uncertainties related to validated trial outcomes. The development and validation of disease-specific, sensitive, robust, and relevant clinical, patient-reported or observer-reported outcomes (PROs; ObsROs) is more challenging in RDs compared to common diseases owing to the rarity and heterogeneity of RDs [54,55]. A recent report suggests solutions for common challenges in PRO and ObsRO assessments in RD clinical trials [56]. It has also been proposed that biomarkers, currently utilized as diagnostic, predictive, or pharmacodynamic tools and surrogate outcomes, may be validated for use as primary endpoints in clinical trials to facilitate the development of novel therapeutics for RDs currently without effective therapies [57].

Patient recruitment for clinical research
Disease rarity poses a hurdle for patient recruitment to any form of research [58,59]. In turn, small study populations complicate the accrual of robust evidence on which to base treatment decisions and HTA evaluations.

Geographic limitations in patient recruitment. Geographic variations in existing clinical expertise [53,60] and disease prevalence influence the ability to recognize and study a RD across geographies [61]. This may result in geographical variability in both the recruitment of sufficient patient numbers for clinical research and the existence of clinical expert centers with the required research capability. Expert centers, such as the European Reference Networks (ERNs), and patient advocacy groups may represent useful sources for patient recruitment for RD research. In 2017, 24 ERNs were launched with the aim of improving diagnosis and treatment of rare or low-prevalence complex diseases or conditions [62]. These virtual networks of RD expert centers, medical specialists, and health care providers share knowledge, best-practice, and some resources for research and the joint delivery of highly specialized health care across EU borders [63].

Insufficiently specific coding systems. Common diagnostic coding systems, such as ICD-9 (International Classification of Diseases, version 9) and ICD-10, may not be sufficiently specific or sensitive for certain RDs [64,65]. A 2015 audit of 6,954 clinical entities listed by Orphanet revealed that only 355 entities had a unique and specific code in ICD-10, and only 162 could be specifically mapped to a set of ICD-10 codes [65]. Patients with an RD may therefore be hard to identify from electronic medical records or other (e.g., administrative) databases. A more rigorous and resource-consuming search for patients with a particular RD may be needed by using, for example, encoded free-text fields. Text mining with artificial intelligence or machine learning is, for the most part, in early stages of development for validation of diagnosis, and together with issues of disease rarity and heterogeneity, complicates epidemiologic research into many RDs.

Orphanet suggested a classification system that endorses representation of 5,400 RDs in the foundation layer of ICD-11 [65], which is 10 times more than in ICD-10 coding. ICD-11 is due to be launched by the World Health Organization (WHO) in 2018, and the adoption of the proposed, more specific, classification of RDs is likely to expedite patient identification and clinical research in RDs.

Ethical and legal hurdles. The use of traditional randomized, controlled trial designs may, in some instances, be deemed inappropriate for interventional studies in RDs; for example, randomization and inclusion of control arms may be unethical [58,59,66,67]. Achieving a sufficient sample size in prospective studies is also hampered by ethical and legal hurdles [59]. The
Marketing authorization may not fulfill the requirements of HTA bodies [72]. The level of scientific and clinical evidence that is deemed sufficient by regulators to grant a novel health technology a marketing authorization may not fulfill the requirements of HTA agencies for issuing a positive reimbursement recommendation [73]. In addition, differences exist in decision-making criteria between different HTA bodies, often resulting in geographic differences in reimbursement status and patient access to RD treatments [6,74].

HTA agencies are primarily concerned with the satisfactory demonstration of a treatment’s value for money, and by the high per patient costs of many RD treatments. Manufacturers generally explain these high costs with the necessity to recover investments made into the development of new RD interventions from small patient populations. Whether such high costs are justified continues to be debated, but there is no doubt that the evaluation of a treatment’s value poses multiple methodological challenges that are particularly pertinent in RDs. The identified challenges relating to the HTA of RD technologies were grouped into two categories, as illustrated in Figure 1.

Nonrandomized trials constituted more than 35% of all reviewed trials, while a control arm was absent in more than 30%, and QoL measures were available in only 27% of trials [78]. Patient registries for RDs are increasingly utilized to gather data on the natural history of a disease, identify patient subgroups, recruit patients for clinical research, and facilitate long-term patient follow-up and real-world data generation [79–83]. Although data from patient registries have, in principle, the potential to address some of the data gaps, their use in standard HTAs still suffers from shortcomings, such as inadequate or inconsistent data collection [13].

No established standard of care. Standard HTA methods are based on a comparative analysis of the treatment under consideration versus the best available treatment alternative. In the absence thereof, an established standard of best supportive care is used as the comparator. Supportive care may differ across geographical regions, and its effectiveness and cost may not be well studied. Therefore, an established standard of care in the management of RDs is often lacking. In addition, routine clinical practice in RD management may differ from that recognized by health authorities; many existing molecules have been used “off-label” in clinical practice to treat RDs where no effective licensed alternatives exist [84–88].

Insufficient knowledge of the natural history of the disease. The shortage of reliable information on the clinical, humanistic, and economic burden of RDs poses a challenge for accurate assessment of the value and impact of a new RD technology. The frequently progressive and degenerative nature of RDs, paired with a poor understanding of the disease’s natural history, is problematic for HTA modeling and projection of long-term treatment outcomes and associated costs [89].

Lack of validated instruments to assess efficacy and effectiveness end points. The absence of robust long-term outcomes data, such as mortality, and validated QoL instruments commonly impedes the reliable estimation of a RD technology’s treatment benefit and outcomes such as the incremental quality-adjusted life years (QALYs), a key metric in cost-effectiveness analysis for many HTA agencies [90]. The development and validation of new, disease-specific and sensitive end points for small RD patient populations can be very time consuming and costly. Approaches for adaptation of existing outcomes for RD populations have been suggested [56], and various disease-specific QoL measures have been developed, such as the Huntington Quality of Life Instrument [91] and the Quality of Life-Primary Ciliary Dyskinesia instrument [92]. The use of biomarkers as primary end points in RD trials has also been proposed [57].

Application of ICER thresholds. When HTA agencies assess RD technologies based on the criterion of cost-effectiveness alone, the resulting ICER is likely to exceed the explicit or implicit threshold value in many jurisdictions [6]. Appraisals driven primarily by cost-effectiveness offer only a partial evaluation of the resulting ICER is frequently well in excess of the “accepted” levels, and the treatment would not be reimbursable according to conventional cost-effectiveness criteria [6]. Whereas some argue that this is not a reason for adopting tailored appraisal methods for RD technologies that grant a premium for rarity [75], others endorse a more holistic appraisal of the value of RD treatments [76,77].

Lack of sufficient and robust clinical data. Owing to the research-related challenges described previously, the available clinical evidence is often limited in RDs. A review of the design of pivotal clinical trials for 64 orphan medicinal products with EMA marketing authorization revealed multiple methodological shortcomings.

Challenges in Health Technology Assessment

Although regulatory approval processes make concessions because of the rarity of certain diseases, reimbursement policy frequently does not follow suit. The European Medicines Agency (EMA) commented in 2013 that a number of new medicines authorized by the EMA are either not reimbursed by national health systems or are not used as expected because they do not meet the requirements of HTA bodies [72].

Standard HTA methods require robust information on comparable efficacy, effectiveness, and associated costs of new health care interventions, including data on morbidity, mortality, quality of life (QoL), and health care utilization. Owing to the challenges of research in RDs and their treatments, not all these types of data may be available at the time of the HTA. HTA methods that are based primarily on cost-effectiveness analyses generally lack sufficient flexibility to allow a comprehensive evaluation that takes account of the complications associated with generating a robust evidence base early in a health intervention’s life cycle. However, this issue is particularly pertinent in RD treatments. Even if a reliable estimation of a RD treatment’s cost-effectiveness can be obtained, its incremental cost-effectiveness ratio (ICER) is frequently well in excess of the “accepted” levels, and the treatment would not be reimbursable according to conventional cost-effectiveness criteria [6]. Whereas some argue that this is not a reason for adopting tailored appraisal methods for RD technologies that grant a premium for rarity [75], others endorse a more holistic appraisal of the value of RD treatments [76,77].

With RDs, it is therefore particularly important to identify and weigh all factors of the health technology that provide value and incorporate them into HTA methods, such as by using multi-criteria decision analysis (MCDA) methodology [96,97]. However, in the absence of MCDA methods specific for RD HTA it is unclear how value elements of a technology, which cannot be captured in a cost-effectiveness evaluation, can be integrated into existing decision-making frameworks to inform price and reimbursement decisions. Although some simulations exist [98,99], it is not known how the incorporation of broader value elements into a decision-making framework would affect the decisions.
HTA agencies are taking steps to develop more tailored processes for the appraisal of RD technologies, in consultation with patients, clinicians, manufacturers, payers, and policymakers. Some HTA bodies apply specific criteria to increase the cost-effectiveness thresholds and thereby facilitate reimbursement of RD technologies. QALY weighting based on disease prevalence was proposed as a way to adapt the standard cost-effectiveness approach for RD interventions [89,100]. In Sweden, the cost-effectiveness threshold is adaptable based on disease severity [74]. The Scottish Medicines Consortium (SMC) may accept a higher than usual ICER for RD treatments if the new medicine meets certain “decision-modifiers,” such as evidence of substantial improvement in life expectancy, QoL, or the absence of alternative therapeutic options of proven benefit [101]. The National Institute for Health and Care Excellence (NICE) established the Highly Specialized Technologies (HST) program for the evaluation of innovative technologies for ultra-rare conditions in England. NICE-HST takes much broader criteria into consideration than NICE’s standard appraisal process and applies significantly higher ICER threshold ranges, dependent on the incremental QALY gain that the new intervention delivers [102]. Germany’s Gemeinsamer Bundesausschuss (G-BA) appraises orphan drugs through a simplified evaluation process that assumes a default additional clinical benefit by virtue of orphan designation, provided their annual sales do not exceed €50 million during the first 12 months of commercialization [103].

Uncertainty for HTA authorities

As a consequence of the research-related hurdles described in the first part of this report, health technology assessors face uncertainty when (1) translating clinical efficacy data from trials into estimates of clinical effectiveness in a real-world setting; (2) evaluating the overall added-value of the new health technology and the extent to which it addresses current unmet medical needs; and (3) quantifying health care costs, utilization, and possible savings over the life time of the disease or patient, in the context of their respective health care setting. Low disease prevalence was found to be associated with a less robust evidence base in HTA submissions for RD interventions [104]. This may leave HTA agencies and health care payers short of the evidence required to make informed and definitive decisions on the reimbursement of an intervention at a given cost. The level of uncertainty in relation to the available evidence base for an RD technology is therefore likely to negatively correlate with its reimbursed price and reimbursement status [105].

Budget constraints, paired with the uncertainties related to effectiveness, value-for-money, and budget impact of RD technologies, increase HTA authorities’ hesitancy in issuing positive reimbursement decisions. Although the impact of RD treatments is relatively minor in terms of proportional health care, and even pharmaceutical, spending [106–109], both total and proportional expenditures on RD treatments are continually rising as a result of the steady growth in the number of RD treatments that are developed and commercialized [110–112].

Concerns over long-term affordability of RD treatments and the limitations of evidence generation in many RDs through conventional clinical trials have compelled policymakers to consider alternative, or complementary, methods for the evaluation of clinical efficacy and cost-effectiveness of new interventions. Such policies usually grant controlled and/or restricted patient access to a new intervention with the expectation that real-world evidence will be generated for the treatment in a clinical setting. These data are intended to close data gaps and enable continuous evaluation of a treatment’s real-world effectiveness, cost-effectiveness, and value. Managed access agreements in England or temporary reimbursement under the G-BA’s orphan drug framework that impose requirements for real-world evidence collection are examples of such programs [103,113]. Although these have methodological challenges and can be resource intensive, they offer a pragmatic solution in the absence of a robust evidence base available at the time of an HTA.

Conclusion

The difficulties faced by different stakeholder groups in RD clinical research and HTA of RD treatments significantly impact the development and the evaluation of these technologies. Research-related challenges, linked predominantly to the low prevalence of RDs, and hurdles related to the HTA of RD technologies result in uncertainty for decision makers in health care. This may consequently impact reimbursement, adoption, and equity of patient access to RD treatments.

The interrelated nature of these challenges requires a collective approach by clinical researchers, industry, HTA agencies, patients, and policymakers toward developing implementable and sustainable solutions. A collaborative effort across stakeholder groups is required to reduce the time and cost for the development of safe and effective new therapies for RDs that address unmet patient needs, provide value for money, and facilitate equitable patient access. Adopting current best practice and developing new approaches are necessary to overcome these challenges and advance RD treatment options and patient access and improve health outcomes. Many organizations, including ISPOR, EURORDIS, the National Organization for Rare Diseases (NORD), the European Network for HTA (EUnetHTA), and the Rare Diseases Clinical Research Network, have developed proposals to address many of the issues highlighted in this report, and which will provide guidance for future RD research.

Acknowledgments

The individual contributions by Matthew Magestro, Lorenzo Man-tovani, Lesley G. Shane, and Vlad Zah are gratefully acknowledged. The authors express special thanks to Ségolène Aymé, Emeritus Director of Research, Founder of Orphanet, Chair of the Topic Advisory Group on Rare Diseases at WHO, Member of the European Commission Expert Group on Rare Diseases, Project Leader of Support IRDiRC, and Editor-in-Chief, Orphanet Journal of Rare Diseases.

The authors also thanks the members of the Challenges in Assessment and Appraisal of Rare Disease Diagnostics & Treatments Working Group, who contributed to the development of this article and have participated as reviewers: Joe Biskupiak, Jacqueline Bowman-Busato, Dyfrig Hughes, Mohit Jain, Katarzyna Kolasa, Zhimei (Jamae) Liu, Phil Ruff, Peter Sun, and Art Zbrozek. Finally, the authors thank the following ISPOR members for valuable feedback on drafts of this report: Alexander Avertyomenko, Caroline Conti, Lincy Lal, Jade Marshall, Gabriel Tremblay, Anup Raj Upreti.

All comments, many of which were substantive and constructive, were considered. ISPOR member comments contribute to the high-quality, consensus nature articles that characterize ISPOR Special Interest Group and Task Force publications. The steady and capable support of ISPOR and its staff is appreciated.

The authors have no other financial relationships to disclose.

REFERENCES


[31] Simoens S. Pricing and reimbursement of orphan drugs: the need for more transparency. Orphanet J Rare Dis 2011;6:42.


Ayme S, Bellet B, Rath A. Rare diseases in ICD11: making rare diseases visible in health information systems through appropriate coding. Orphanet J Rare Dis 2015;10:35.


